



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 16, 2017 – 01:01 AM EST

PDB ID : 5T1P  
Title : Crystal structure of the putative periplasmic solute-binding protein from *Campylobacter jejuni*  
Authors : Filippova, E.V.; Wawrzak, Z.; Sandoval, J.; Skarina, T.; Grimshaw, S.; Savchenko, A.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : unknown  
Resolution : 2.00 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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-20030345

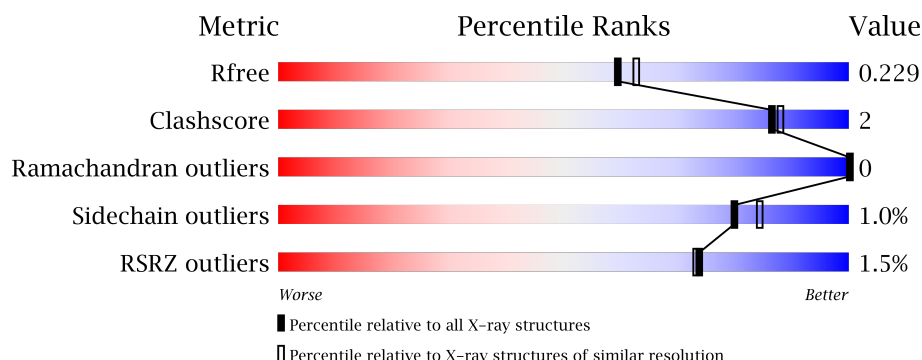
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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	348	<div> <div>0.1%</div> <div>86% 7% 6%</div> </div>
1	B	348	<div> <div>87% 6% 6%</div> </div>
1	C	348	<div> <div>4%</div> <div>88% 5% 6%</div> </div>
1	D	348	<div> <div>3%</div> <div>90% 5% 6%</div> </div>
1	E	348	<div> <div>89% 5% 6%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	348	 .% 89% 5% 6%
1	G	348	 85% 8% 6%
1	H	348	 2% 86% 6% 6%

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	PEG	A	401	-	-	-	X
2	PEG	F	402	-	-	-	X
4	BTB	E	401	-	-	-	X
4	BTB	G	401	-	-	-	X
5	TRS	F	401	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 22353 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 ABC transporter, periplasmic substrate-binding protein.

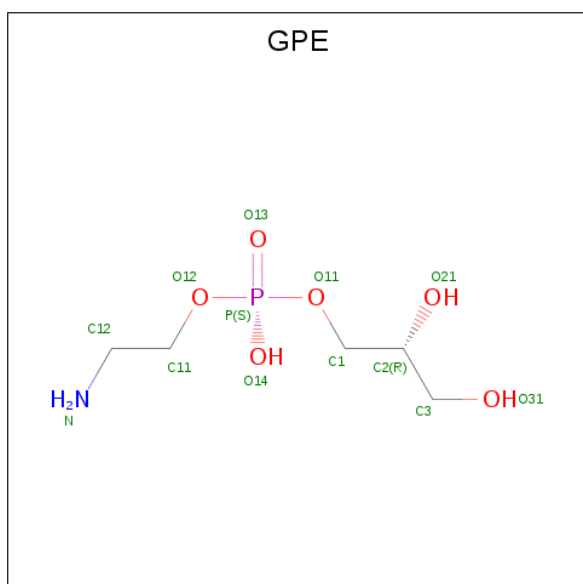
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	327	Total	C	N	O	Se	0	6	0
			2607	1663	442	498	4			
1	B	326	Total	C	N	O	Se	0	3	0
			2573	1642	437	491	3			
1	C	326	Total	C	N	O	Se	0	1	0
			2557	1631	433	490	3			
1	D	328	Total	C	N	O	Se	0	0	0
			2562	1634	434	491	3			
1	E	326	Total	C	N	O	Se	0	3	0
			2576	1642	438	493	3			
1	F	327	Total	C	N	O	Se	0	2	0
			2568	1637	434	494	3			
1	G	326	Total	C	N	O	Se	0	3	0
			2577	1642	438	494	3			
1	H	326	Total	C	N	O	Se	0	6	0
			2595	1653	441	496	5			

- Molecule 2 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



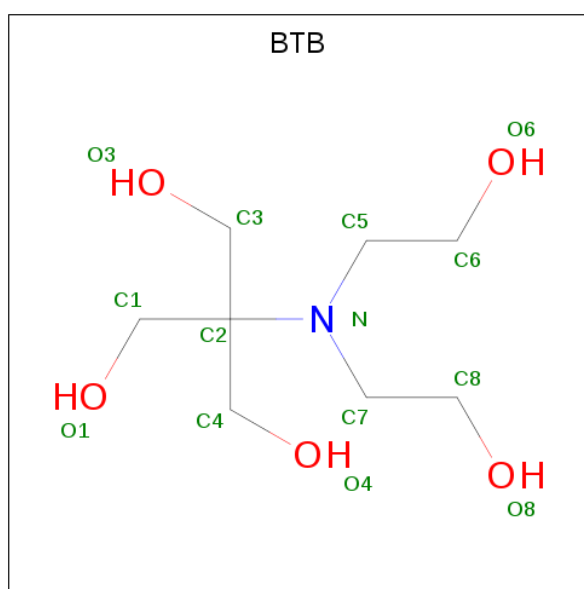
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			7	4	3		
2	E	1	Total	C	O	0	0
			4	2	2		
2	F	1	Total	C	O	0	0
			4	2	2		
2	H	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is L-ALPHA-GLYCEROPHOSPHORYLETHANOLAMINE (three-letter code: GPE) (formula:  $C_5H_{14}NO_6P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			13	5	1	6	1		
3	B	1	Total	C	N	O	P	0	0
			13	5	1	6	1		
3	C	1	Total	C	N	O	P	0	0
			13	5	1	6	1		
3	D	1	Total	C	N	O	P	0	0
			13	5	1	6	1		

- Molecule 4 is 2-[BIS-(2-HYDROXY-ETHYL)-AMINO]-2-HYDROXYMETHYL-PROPAN E-1,3-DIOL (three-letter code: BTB) (formula: C<sub>8</sub>H<sub>19</sub>NO<sub>5</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	F	1	Total	C	N	O	0	0
			8	4	1	3		
5	H	1	Total	C	N	O	0	0
			8	4	1	3		

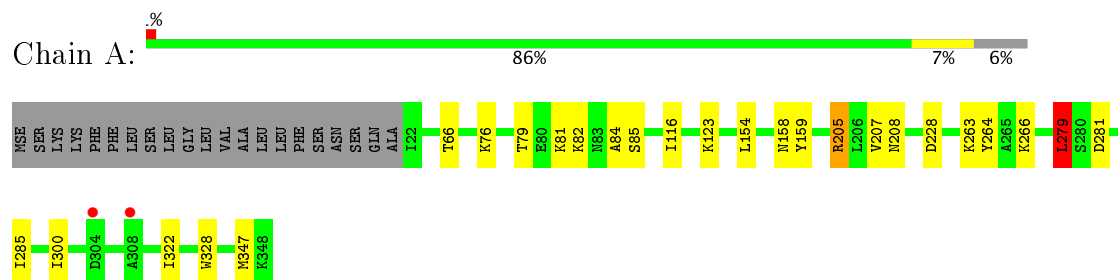
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	267	Total	O	0	2
			269	269		
6	B	208	Total	O	0	0
			208	208		
6	C	145	Total	O	0	0
			145	145		
6	D	219	Total	O	0	0
			219	219		
6	E	254	Total	O	0	2
			255	255		
6	F	192	Total	O	0	0
			192	192		
6	G	189	Total	O	0	0
			189	189		
6	H	145	Total	O	0	1
			146	146		

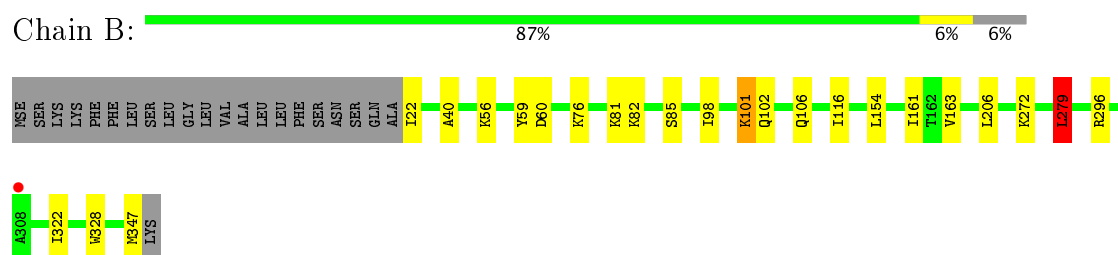
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

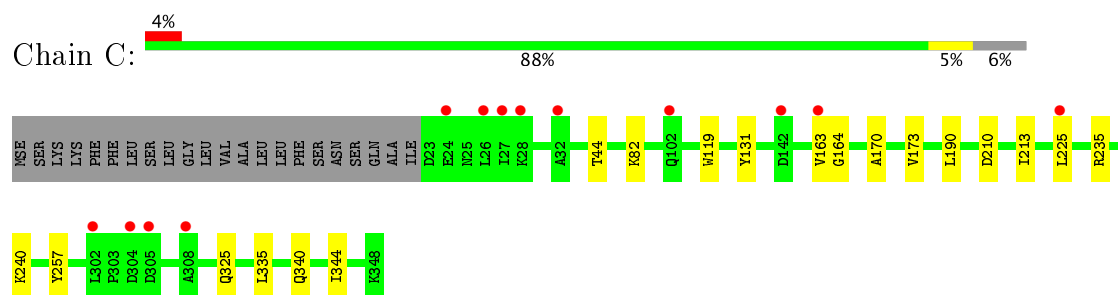
- Molecule 1: ABC transporter, periplasmic substrate-binding protein



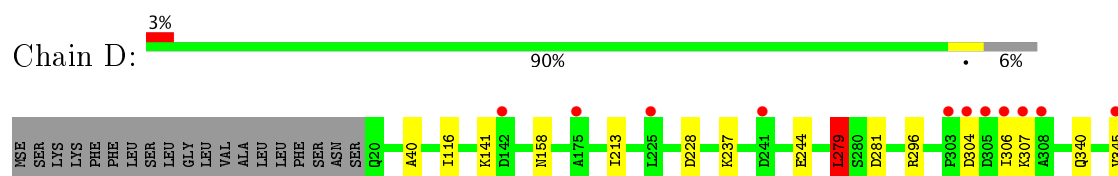
- Molecule 1: ABC transporter, periplasmic substrate-binding protein



- Molecule 1: ABC transporter, periplasmic substrate-binding protein




- Molecule 1: ABC transporter, periplasmic substrate-binding protein






D346  
K347  
LYS

- Molecule 1: ABC transporter, periplasmic substrate-binding protein

Chain E:  89% 5% 6%


MSE SER LYS PHE LEU SER LEU GLY VAL ALA LEU PHE SER ASN SER GLN ALA I22 Y59 D60 I61 L128 K141 V144 K145 N208 D236 E244 K289 R293 R296 P303 I306 S313 E334 K347 LYS

- Molecule 1: ABC transporter, periplasmic substrate-binding protein

Chain F:  89% 5% 6%

MSE SER LYS PHE LEU SER LEU GLY VAL ALA LEU PHE SER ASN SER GLN A21 E24 Y36 F77 G89 Q102 S103 V104 L128 K156 Y168 I213 T259 A270 R293 P312 S313 E314 K326 K347 LYS


- Molecule 1: ABC transporter, periplasmic substrate-binding protein

Chain G:  85% 8% 6%

MSE SER LYS PHE LEU SER LEU GLY VAL ALA LEU PHE SER ASN SER GLN I22 V36 I74 F95 K101 Q102 L128 L129 V143 V144 Y159 V166 I213 R205 T213 A214 I215 K218 V240 D250 G256 A270 Y291 A292

R293 I297 L302 P303 I306 K309 S313 E314 W339 V343 K347 LYS

- Molecule 1: ABC transporter, periplasmic substrate-binding protein

Chain H:  86% 2% 6% 6%

MSE SER LYS PHE LEU SER LEU GLY VAL ALA LEU PHE SER ASN SER GLN A21 L26 I27 K28 L39 A40 M41 M68 Q72 E73 K76 K82 ASH A84 S85 I88 G88 D90 V91 F95 K101 V104 T110 A120 L128 L129

R205 Y257 T258 T259 N262 A265 K266 I278 L279 S280 D281 I285 R293 R296 L302 K347 LYS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.19Å 89.28Å 100.39Å 68.88° 82.52° 70.66°	Depositor
Resolution (Å)	30.00 – 2.00 29.67 – 2.00	Depositor EDS
% Data completeness (in resolution range)	95.8 (30.00-2.00) 85.0 (29.67-2.00)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.18 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, $R_{free}$	0.176 , 0.224 0.184 , 0.229	Depositor DCC
$R_{free}$ test set	8256 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.4	Xtriage
Anisotropy	0.105	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 43.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	22353	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 48.30 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 8.7704e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: GPE, TRS, PEG, BTB

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	0.67	0/2656	0.87	4/3591 (0.1%)
1	B	0.62	0/2622	0.85	4/3550 (0.1%)
1	C	0.60	0/2606	0.81	0/3526
1	D	0.60	0/2611	0.86	5/3535 (0.1%)
1	E	0.67	1/2625 (0.0%)	0.88	3/3553 (0.1%)
1	F	0.58	0/2617	0.81	0/3543
1	G	0.61	0/2626	0.82	2/3555 (0.1%)
1	H	0.57	0/2643	0.83	5/3575 (0.1%)
All	All	0.62	1/21006 (0.0%)	0.84	23/28428 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	334	GLU	CD-OE1	5.47	1.31	1.25

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	296	ARG	NE-CZ-NH2	-10.83	114.89	120.30
1	E	208	ASN	CB-CA-C	-10.80	88.80	110.40
1	D	296	ARG	NE-CZ-NH1	9.17	124.88	120.30
1	A	205	ARG	NE-CZ-NH1	7.82	124.21	120.30
1	A	205	ARG	NE-CZ-NH2	-7.57	116.52	120.30
1	D	279	LEU	CA-CB-CG	-6.76	99.76	115.30
1	B	347	MSE	CG-SE-CE	-6.48	84.64	98.90
1	G	205	ARG	NE-CZ-NH1	5.94	123.27	120.30
1	A	279	LEU	CA-CB-CG	-5.93	101.66	115.30
1	H	296	ARG	NE-CZ-NH2	-5.88	117.36	120.30
1	B	279	LEU	CA-CB-CG	-5.78	102.00	115.30
1	D	281	ASP	CB-CG-OD2	-5.72	113.15	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	296	ARG	NE-CZ-NH1	5.65	123.13	120.30
1	G	250	ASP	CB-CG-OD1	5.61	123.35	118.30
1	H	296	ARG	NE-CZ-NH1	5.40	123.00	120.30
1	H	68[A]	MSE	CG-SE-CE	-5.38	87.07	98.90
1	H	68[B]	MSE	CG-SE-CE	-5.38	87.07	98.90
1	D	228	ASP	CB-CG-OD1	5.25	123.02	118.30
1	E	347	MSE	N-CA-CB	5.18	119.93	110.60
1	A	228	ASP	CB-CG-OD1	5.09	122.88	118.30
1	B	296	ARG	NE-CZ-NH2	-5.09	117.76	120.30
1	H	205	ARG	NE-CZ-NH1	5.08	122.84	120.30
1	B	101	LYS	CA-CB-CG	5.04	124.48	113.40

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	2607	0	2613	17	0
1	B	2573	0	2575	15	0
1	C	2557	0	2557	14	0
1	D	2562	0	2560	7	0
1	E	2576	0	2575	8	0
1	F	2568	0	2561	9	0
1	G	2577	0	2572	15	0
1	H	2595	0	2588	19	0
2	A	7	0	10	0	0
2	E	4	0	5	0	0
2	F	4	0	5	0	0
2	H	4	0	5	0	0
3	A	13	0	13	0	0
3	B	13	0	13	1	0
3	C	13	0	13	0	0
3	D	13	0	13	1	0
4	E	14	0	19	0	0
4	G	14	0	19	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	F	8	0	12	0	0
5	H	8	0	12	0	0
6	A	269	0	0	2	0
6	B	208	0	0	1	0
6	C	145	0	0	2	0
6	D	219	0	0	0	0
6	E	255	0	0	0	0
6	F	192	0	0	1	0
6	G	189	0	0	1	0
6	H	146	0	0	0	0
All	All	22353	0	20740	102	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 2.

All (102) 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:H:293[A]:ARG:HH11	1:H:293[A]:ARG:HG3	1.16	1.02
1:H:91:VAL:CG1	1:H:95:PHE:HB2	2.04	0.88
1:H:293[A]:ARG:NH1	1:H:293[A]:ARG:HG3	1.84	0.81
1:H:281:ASP:O	1:H:285:ILE:HG12	1.83	0.78
1:H:91:VAL:HG13	1:H:95:PHE:HB2	1.69	0.74
1:B:154:LEU:HD21	1:B:161:ILE:HD13	1.69	0.74
1:H:293[A]:ARG:NH1	1:H:293[A]:ARG:CG	2.58	0.67
1:H:41[A]:MSE:HE1	1:H:88:ILE:HG12	1.78	0.64
1:A:263:LYS:HE2	1:A:264:TYR:CZ	2.33	0.64
1:D:306:ILE:HD12	1:D:307:LYS:N	2.15	0.61
1:C:163[B]:VAL:HG13	1:C:225:LEU:HD12	1.83	0.60
1:C:119:TRP:O	1:C:325:GLN:NE2	2.35	0.59
1:B:154:LEU:HD23	1:B:154:LEU:C	2.22	0.59
1:G:293[B]:ARG:NH1	6:G:505:HOH:O	2.35	0.59
1:A:84:ALA:O	1:A:266:LYS:HD3	2.01	0.58
1:E:293[A]:ARG:NH2	1:E:313:SER:OG	2.37	0.58
1:B:116:ILE:HD11	1:B:279:LEU:HD22	1.88	0.56
1:G:159:TYR:O	1:G:205:ARG:HD2	2.06	0.55
1:C:340:GLN:NE2	6:C:503:HOH:O	2.36	0.55
1:B:98:ILE:CG2	1:B:102:GLN:HE21	2.20	0.54
1:F:213:ILE:HD12	1:F:213:ILE:H	1.73	0.54
1:G:291:TYR:CE1	1:G:309:LYS:HG3	2.42	0.54
1:A:208:ASN:ND2	1:A:347[B]:MSE:SE	2.91	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:281:ASP:O	1:A:285:ILE:HG12	2.09	0.53
1:G:36:VAL:HB	1:G:270:ALA:HB1	1.90	0.53
1:F:293:ARG:NH2	1:F:313:SER:OG	2.42	0.53
1:A:159:TYR:O	1:A:205:ARG:HD2	2.09	0.52
1:A:76[A]:LYS:HE2	1:A:85:SER:OG	2.10	0.52
1:A:116:ILE:HD11	1:A:279:LEU:HD22	1.92	0.52
1:G:129:LEU:HD13	1:G:256:GLY:HA3	1.91	0.51
1:H:41[A]:MSE:HE3	1:H:89:GLY:C	2.31	0.51
1:E:289:LYS:HG3	1:E:306:ILE:HD13	1.91	0.51
1:C:164:GLY:HA2	1:C:210:ASP:HA	1.93	0.51
1:G:293[A]:ARG:NH2	1:G:313:SER:OG	2.43	0.51
1:E:303:PRO:HG2	1:E:306:ILE:HD12	1.93	0.50
1:F:24:GLU:CD	1:F:24:GLU:H	2.14	0.50
1:H:110:THR:HG21	1:H:279:LEU:HD12	1.94	0.50
1:H:285:ILE:HD12	1:H:302:LEU:HD23	1.93	0.50
1:D:213:ILE:HD13	1:D:237:LYS:HD3	1.94	0.50
1:G:339:TRP:O	1:G:343:VAL:HG22	2.12	0.50
1:H:128:LEU:HD12	1:H:128:LEU:O	2.11	0.49
1:A:79:THR:HG23	1:C:344:ILE:HD13	1.95	0.49
1:B:322:ILE:HD13	1:B:328:TRP:CD1	2.47	0.49
1:H:285:ILE:CD1	1:H:302:LEU:HD23	2.43	0.49
1:C:190:LEU:HD12	1:C:335:LEU:HD13	1.94	0.48
1:A:84:ALA:O	1:A:266:LYS:CD	2.61	0.48
1:D:141:LYS:HD3	1:D:244:GLU:HB2	1.95	0.48
1:H:91:VAL:CG1	1:H:95:PHE:CB	2.85	0.47
1:B:163[A]:VAL:HG21	1:B:206:LEU:CD1	2.43	0.47
1:F:89:GLY:O	1:F:259:THR:HA	2.15	0.47
1:E:141:LYS:HD3	1:E:244:GLU:HB2	1.97	0.47
1:E:59:TYR:HB2	1:E:61:ILE:HD13	1.97	0.47
1:H:120:ALA:HB1	1:H:129:LEU:HB2	1.97	0.47
1:F:77:PHE:HB2	1:F:104:VAL:HG13	1.96	0.47
1:C:235:ARG:NH2	6:C:501:HOH:O	2.30	0.47
1:A:207:VAL:HG22	1:C:82:LYS:HG3	1.96	0.47
1:C:163[B]:VAL:O	1:C:163[B]:VAL:HG12	2.13	0.47
1:D:40:ALA:HB1	3:D:401:GPE:H121	1.97	0.47
1:B:98:ILE:HG23	1:B:102:GLN:HE21	1.81	0.46
1:A:285:ILE:HD13	1:A:300:ILE:HD12	1.97	0.46
1:A:322:ILE:HD13	1:A:328:TRP:CD1	2.50	0.45
1:A:66[B]:THR:HG22	6:A:717:HOH:O	2.15	0.45
1:B:56:LYS:O	1:B:60:ASP:HA	2.16	0.45
1:C:131:TYR:CZ	1:C:257:TYR:HB2	2.52	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:190:LEU:CD1	1:C:335:LEU:HD13	2.47	0.45
1:C:235:ARG:HG2	1:C:240:LYS:HA	1.98	0.45
1:G:166:VAL:HG21	1:G:339:TRP:CG	2.51	0.45
1:F:77:PHE:CB	1:F:104:VAL:HG13	2.47	0.44
1:H:76:LYS:HE2	1:H:85:SER:OG	2.17	0.44
1:F:312:PRO:HB2	1:F:314[B]:GLU:HG3	1.99	0.44
1:F:102:GLN:NE2	6:F:508:HOH:O	2.50	0.44
1:B:76:LYS:HE3	1:B:85:SER:OG	2.18	0.44
1:E:303:PRO:CG	1:E:306:ILE:HD12	2.48	0.44
1:B:106:GLN:OE1	1:B:272:LYS:NZ	2.51	0.43
1:H:262:ASN:HB3	1:H:265:ALA:HB2	2.00	0.43
1:H:68[B]:MSE:CE	1:H:73:GLU:HA	2.48	0.43
1:A:154:LEU:O	1:A:205:ARG:NH2	2.49	0.43
1:H:68[A]:MSE:SE	1:H:72:GLN:HB3	2.69	0.42
1:A:81:LYS:HB3	1:A:82:LYS:H	1.74	0.42
1:E:145:LYS:HD2	1:E:145:LYS:N	2.34	0.42
1:F:36:VAL:HB	1:F:270:ALA:HB1	1.99	0.42
1:G:213:ILE:HD12	1:G:214:ALA:H	1.85	0.42
1:B:22:ILE:HD12	1:B:59:TYR:OH	2.18	0.42
1:A:158:ASN:HA	1:A:205:ARG:HD3	2.02	0.42
1:G:101:LYS:HE3	1:G:102:GLN:HE21	1.85	0.42
1:B:40:ALA:HB1	3:B:401:GPE:H121	2.01	0.41
1:C:170:ALA:HA	1:C:173:VAL:HG22	2.02	0.41
1:A:66[B]:THR:HG23	6:A:537:HOH:O	2.20	0.41
1:C:44:THR:HG21	1:C:213:ILE:HD12	2.02	0.41
1:G:215:ASN:OD1	1:G:218:LYS:NZ	2.50	0.41
1:B:163[A]:VAL:HG21	1:B:206:LEU:HD11	2.02	0.41
1:G:74:ILE:HD11	1:G:95:PHE:HB3	2.03	0.41
1:B:22:ILE:HG12	6:B:510:HOH:O	2.21	0.41
1:D:116:ILE:HD11	1:D:279:LEU:HD22	2.02	0.41
1:D:158:ASN:OD1	1:D:158:ASN:N	2.54	0.41
1:G:303:PRO:HD2	1:G:306:ILE:HD12	2.02	0.41
1:H:257:TYR:N	1:H:257:TYR:CD1	2.89	0.40
1:B:81:LYS:HB3	1:B:82:LYS:H	1.69	0.40
1:D:340:GLN:O	1:D:345:VAL:HG13	2.22	0.40
1:G:297:ILE:CD1	1:G:302:LEU:HD11	2.52	0.40
1:E:59:TYR:HB2	1:E:61:ILE:CD1	2.52	0.40
1:G:143:VAL:HG23	1:G:144:VAL:HG23	2.03	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	331/348 (95%)	321 (97%)	10 (3%)	0	100	100
1	B	327/348 (94%)	319 (98%)	8 (2%)	0	100	100
1	C	325/348 (93%)	313 (96%)	12 (4%)	0	100	100
1	D	326/348 (94%)	317 (97%)	9 (3%)	0	100	100
1	E	327/348 (94%)	317 (97%)	10 (3%)	0	100	100
1	F	327/348 (94%)	319 (98%)	8 (2%)	0	100	100
1	G	327/348 (94%)	317 (97%)	10 (3%)	0	100	100
1	H	328/348 (94%)	318 (97%)	10 (3%)	0	100	100
All	All	2618/2784 (94%)	2541 (97%)	77 (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	273/281 (97%)	271 (99%)	2 (1%)	87	90
1	B	269/281 (96%)	267 (99%)	2 (1%)	87	90
1	C	267/281 (95%)	267 (100%)	0	100	100
1	D	267/281 (95%)	265 (99%)	2 (1%)	87	90
1	E	269/281 (96%)	265 (98%)	4 (2%)	70	74
1	F	268/281 (95%)	265 (99%)	3 (1%)	78	82

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	269/281 (96%)	265 (98%)	4 (2%)	70	74
1	H	270/281 (96%)	265 (98%)	5 (2%)	62	66
All	All	2152/2248 (96%)	2130 (99%)	22 (1%)	80	84

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	123	LYS
1	A	279	LEU
1	B	101	LYS
1	B	279	LEU
1	D	279	LEU
1	D	304	ASP
1	E	128	LEU
1	E	144	VAL
1	E	236	ASP
1	E	347	MSE
1	F	128	LEU
1	F	326	LYS
1	F	347	MSE
1	G	128	LEU
1	G	213	ILE
1	G	240	LYS
1	G	314	GLU
1	H	104	VAL
1	H	128	LEU
1	H	129	LEU
1	H	293[A]	ARG
1	H	293[B]	ARG

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	208	ASN
1	A	325	GLN
1	B	102	GLN
1	E	31	GLN
1	E	215	ASN
1	F	102	GLN
1	F	208	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	209	ASN
1	G	102	GLN
1	G	269	ASN
1	H	31	GLN
1	H	325	GLN

### 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](#)

12 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	PEG	A	401	-	6,6,6	0.53	0	5,5,5	0.34	0
3	GPE	A	402	-	12,12,12	0.53	0	12,15,15	0.45	0
3	GPE	B	401	-	12,12,12	0.35	0	12,15,15	0.56	0
3	GPE	C	401	-	12,12,12	0.44	0	12,15,15	0.44	0
3	GPE	D	401	-	12,12,12	0.40	0	12,15,15	0.53	0
4	BTB	E	401	-	13,13,13	0.86	0	9,16,16	0.89	0
2	PEG	E	402	-	3,3,6	0.44	0	2,2,5	0.28	0
5	TRS	F	401	-	7,7,7	0.50	0	9,9,9	0.41	0
2	PEG	F	402	-	3,3,6	0.43	0	2,2,5	0.27	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	BTB	G	401	-	13,13,13	0.94	0	9,16,16	1.60	2 (22%)
5	TRS	H	401	-	7,7,7	0.42	0	9,9,9	0.55	0
2	PEG	H	402	-	3,3,6	0.40	0	2,2,5	0.36	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	PEG	A	401	-	-	0/4/4/4	0/0/0/0
3	GPE	A	402	-	-	0/13/13/13	0/0/0/0
3	GPE	B	401	-	-	0/13/13/13	0/0/0/0
3	GPE	C	401	-	-	0/13/13/13	0/0/0/0
3	GPE	D	401	-	-	0/13/13/13	0/0/0/0
4	BTB	E	401	-	-	0/21/21/21	0/0/0/0
2	PEG	E	402	-	-	0/1/1/4	0/0/0/0
5	TRS	F	401	-	-	0/9/9/9	0/0/0/0
2	PEG	F	402	-	-	0/1/1/4	0/0/0/0
4	BTB	G	401	-	-	0/21/21/21	0/0/0/0
5	TRS	H	401	-	-	0/9/9/9	0/0/0/0
2	PEG	H	402	-	-	0/1/1/4	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	401	BTB	C7-N-C2	2.06	119.80	113.70
4	G	401	BTB	C8-C7-N	3.17	123.97	111.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	401	GPE	1	0
3	D	401	GPE	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	324/348 (93%)	-0.31	2 (0%) 89 88	10, 20, 40, 58	0
1	B	323/348 (92%)	-0.20	1 (0%) 93 93	11, 25, 44, 58	0
1	C	323/348 (92%)	0.14	13 (4%) 39 39	17, 32, 55, 82	0
1	D	325/348 (93%)	-0.12	11 (3%) 46 46	13, 25, 55, 96	0
1	E	323/348 (92%)	-0.28	1 (0%) 93 93	11, 23, 45, 62	0
1	F	324/348 (93%)	-0.17	2 (0%) 89 88	16, 29, 50, 61	0
1	G	323/348 (92%)	-0.18	1 (0%) 93 93	14, 30, 56, 70	0
1	H	323/348 (92%)	0.09	7 (2%) 62 61	16, 35, 65, 102	0
All	All	2588/2784 (92%)	-0.13	38 (1%) 74 73	10, 27, 54, 102	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	308	ALA	6.3
1	D	307	LYS	6.0
1	D	304	ASP	4.4
1	C	308	ALA	4.3
1	E	208	ASN	4.3
1	H	101	LYS	3.5
1	C	28	LYS	3.4
1	D	306	ILE	3.3
1	C	304	ASP	3.2
1	D	225	LEU	2.8
1	A	308	ALA	2.8
1	D	142	ASP	2.8
1	A	304	ASP	2.8
1	C	24	GLU	2.7
1	H	259	THR	2.7
1	C	102	GLN	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	156	LYS	2.6
1	C	305	ASP	2.5
1	C	163[A]	VAL	2.5
1	C	142	ASP	2.5
1	H	278	ILE	2.5
1	H	28	LYS	2.5
1	D	241	ASP	2.4
1	D	345	VAL	2.4
1	C	26	LEU	2.4
1	B	308	ALA	2.4
1	H	26	LEU	2.3
1	C	32	ALA	2.3
1	C	27	ILE	2.3
1	C	302	LEU	2.3
1	H	266	LYS	2.3
1	H	39	LEU	2.2
1	G	213	ILE	2.2
1	D	305	ASP	2.2
1	D	303	PRO	2.1
1	C	225	LEU	2.1
1	F	168	VAL	2.1
1	D	175	ALA	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	PEG	A	401	7/7	0.81	0.21	5.15	49,54,60,66	0
2	PEG	F	402	4/7	0.90	0.15	3.98	45,46,46,46	0
4	BTB	E	401	14/14	0.88	0.15	2.48	33,36,46,55	0
5	TRS	F	401	8/8	0.76	0.20	2.36	42,49,54,54	0
4	BTB	G	401	14/14	0.83	0.15	2.34	27,30,32,34	0
5	TRS	H	401	8/8	0.73	0.20	1.90	53,58,61,62	0
2	PEG	E	402	4/7	0.92	0.13	1.33	44,46,48,52	0
3	GPE	D	401	13/13	0.97	0.17	1.32	15,19,30,35	0
3	GPE	B	401	13/13	0.98	0.13	0.68	13,14,26,28	0
3	GPE	C	401	13/13	0.97	0.15	0.42	20,22,30,32	0
3	GPE	A	402	13/13	0.99	0.10	-0.04	11,12,23,27	0
2	PEG	H	402	4/7	0.84	0.20	-	56,56,59,59	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.