



# Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 03:35 pm BST

PDB ID : 3PFV  
Title : Crystal structure of Cbl-b TKB domain in complex with EGFR pY1069 peptide  
Authors : Chaikuad, A.; Guo, K.; Cooper, C.D.O.; Ayinampudi, V.; Krojer, T.; Muniz, J.R.C.; Vollmar, M.; Canning, P.; Gileadi, O.; von Delft, F.; Arrowsmith, C.H.; Weigelt, J.; Edwards, A.M.; Bountra, C.; Bullock, A.; Structural Genomics Consortium (SGC)  
Deposited on : 2010-10-29  
Resolution : 2.27 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

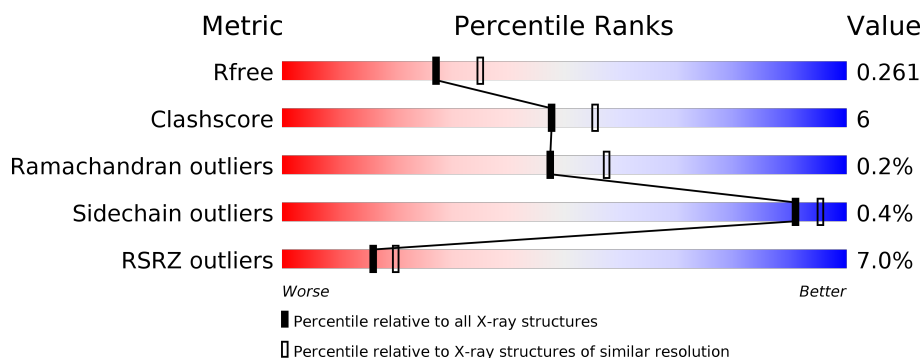
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.27 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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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 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	315	<div> <div>5%</div> <div> <div></div> <div>82%</div> <div>13%</div> <div>• •</div> </div> </div>
1	B	315	<div> <div>8%</div> <div> <div></div> <div>84%</div> <div>13%</div> <div>• •</div> </div> </div>
2	C	11	<div> <div>18%</div> <div> <div></div> <div>82%</div> <div>9%</div> <div>9%</div> </div> </div>
2	D	11	<div> <div>18%</div> <div> <div></div> <div>91%</div> <div>9%</div> </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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	CL	A	3	-	-	X	-

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 5435 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 E3 ubiquitin-protein ligase CBL-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	302	Total	C	N	O	S	0	4	0
			2485	1610	423	442	10			
1	B	310	Total	C	N	O	S	0	3	0
			2507	1627	423	447	10			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	MET	-	INITIATING METHIONINE	UNP Q13191
A	345	ALA	-	EXPRESSION TAG	UNP Q13191
A	346	GLU	-	EXPRESSION TAG	UNP Q13191
A	347	ASN	-	EXPRESSION TAG	UNP Q13191
A	348	LEU	-	EXPRESSION TAG	UNP Q13191
A	349	TYR	-	EXPRESSION TAG	UNP Q13191
A	350	PHE	-	EXPRESSION TAG	UNP Q13191
A	351	GLN	-	EXPRESSION TAG	UNP Q13191
B	37	MET	-	INITIATING METHIONINE	UNP Q13191
B	345	ALA	-	EXPRESSION TAG	UNP Q13191
B	346	GLU	-	EXPRESSION TAG	UNP Q13191
B	347	ASN	-	EXPRESSION TAG	UNP Q13191
B	348	LEU	-	EXPRESSION TAG	UNP Q13191
B	349	TYR	-	EXPRESSION TAG	UNP Q13191
B	350	PHE	-	EXPRESSION TAG	UNP Q13191
B	351	GLN	-	EXPRESSION TAG	UNP Q13191

- Molecule 2 is a protein called 11-meric peptide from Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	10	Total	C	N	O	P	0	0	0
			75	42	13	19	1			
2	D	11	Total	C	N	O	P	0	0	0
			80	45	14	20	1			

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

*Continued on next page...*

Continued from previous page...

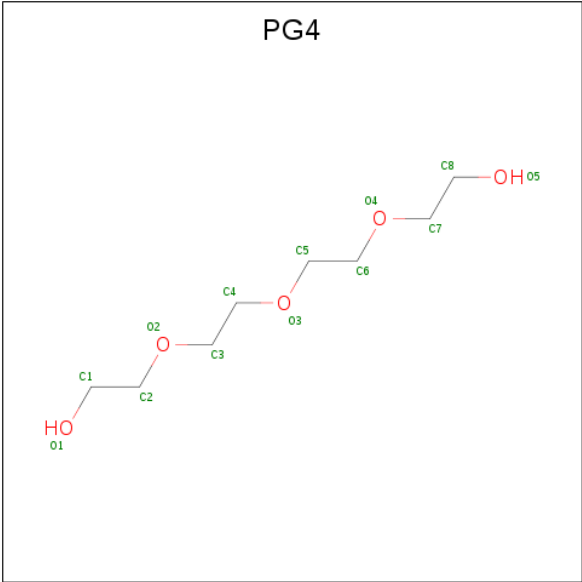
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	1
			8	4	4		
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			13	8	5		

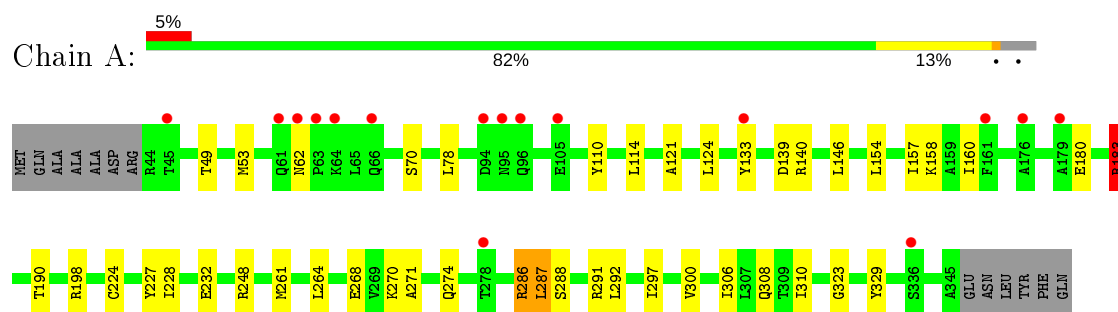
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	107	Total	O	0	0
			107	107		
8	B	100	Total	O	0	0
			100	100		
8	C	5	Total	O	0	0
			5	5		
8	D	8	Total	O	0	0
			8	8		

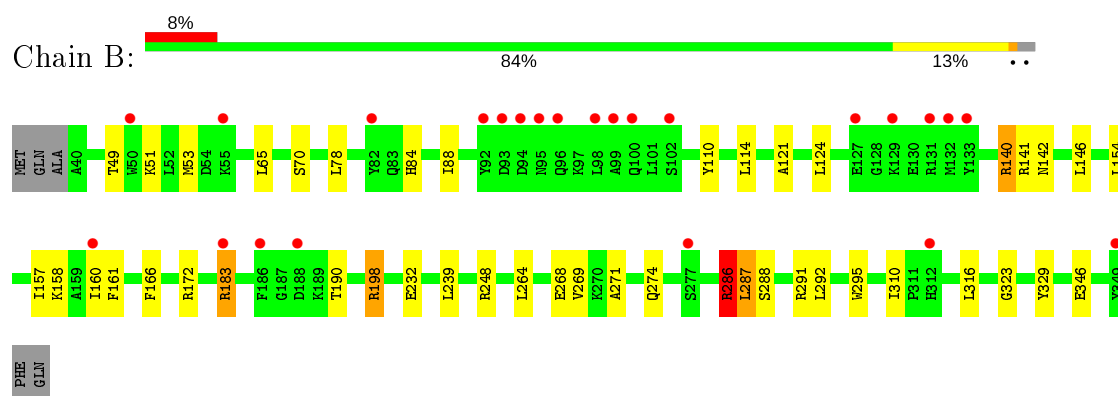
### 3 Residue-property plots [i](#)

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.

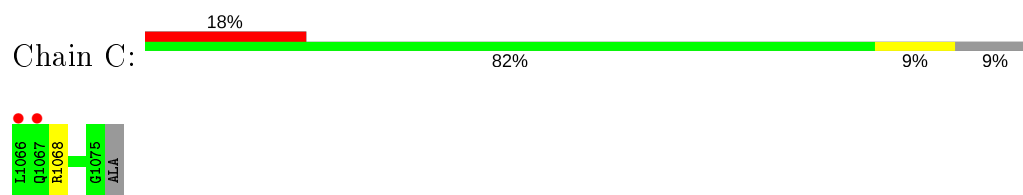
- Molecule 1: E3 ubiquitin-protein ligase CBL-B



- Molecule 1: E3 ubiquitin-protein ligase CBL-B



- Molecule 2: 11-meric peptide from Epidermal growth factor receptor



- Molecule 2: 11-meric peptide from Epidermal growth factor receptor







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.06 Å 98.93 Å 61.98 Å 90.00° 110.63° 90.00°	Depositor
Resolution (Å)	49.47 – 2.27 48.87 – 2.27	Depositor EDS
% Data completeness (in resolution range)	96.1 (49.47-2.27) 96.1 (48.87-2.27)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.95 (at 2.27 Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, $R_{free}$	0.216 , 0.262 0.213 , 0.261	Depositor DCC
$R_{free}$ test set	1523 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.8	Xtriage
Anisotropy	0.164	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 45.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.028 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5435	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.49% 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: CL, NA, EDO, PG4, SO4, PTR

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	5/2561 (0.2%)	1.17	15/3465 (0.4%)
1	B	0.79	4/2581 (0.2%)	1.11	16/3495 (0.5%)
2	C	0.71	0/58	0.92	0/76
2	D	0.60	0/63	0.81	0/83
All	All	0.78	9/5263 (0.2%)	1.14	31/7119 (0.4%)

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	2
All	All	0	3

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	286	ARG	CZ-NH2	-11.47	1.18	1.33
1	A	140	ARG	CZ-NH1	-10.15	1.19	1.33
1	B	140	ARG	CZ-NH1	-9.99	1.20	1.33
1	A	286	ARG	CZ-NH1	-9.98	1.20	1.33
1	A	286	ARG	CZ-NH2	-9.88	1.20	1.33
1	B	286	ARG	CZ-NH1	-8.54	1.22	1.33
1	B	140	ARG	CZ-NH2	-7.23	1.23	1.33
1	A	224	CYS	CB-SG	5.22	1.91	1.82
1	A	140	ARG	CZ-NH2	-5.08	1.26	1.33

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	140	ARG	NE-CZ-NH2	24.28	132.44	120.30
1	A	198	ARG	NE-CZ-NH1	-23.19	108.70	120.30
1	B	183	ARG	NE-CZ-NH1	-21.30	109.65	120.30
1	A	183	ARG	NE-CZ-NH2	-20.82	109.89	120.30
1	B	198	ARG	NE-CZ-NH2	-20.27	110.17	120.30
1	B	286	ARG	NE-CZ-NH1	18.93	129.77	120.30
1	A	198	ARG	NE-CZ-NH2	16.52	128.56	120.30
1	A	183	ARG	NE-CZ-NH1	16.34	128.47	120.30
1	B	183	ARG	NE-CZ-NH2	16.23	128.41	120.30
1	A	286	ARG	NE-CZ-NH1	14.34	127.47	120.30
1	A	286	ARG	NE-CZ-NH2	13.56	127.08	120.30
1	B	198	ARG	NE-CZ-NH1	13.20	126.90	120.30
1	A	286	ARG	NH1-CZ-NH2	-12.70	105.44	119.40
1	B	286	ARG	NH1-CZ-NH2	-12.51	105.64	119.40
1	B	140	ARG	NE-CZ-NH2	11.79	126.19	120.30
1	B	140	ARG	CG-CD-NE	-10.46	89.84	111.80
1	B	140	ARG	NH1-CZ-NH2	-10.19	108.20	119.40
1	A	140	ARG	NH1-CZ-NH2	-9.81	108.60	119.40
1	A	198	ARG	CD-NE-CZ	9.78	137.30	123.60
1	B	140	ARG	NE-CZ-NH1	9.13	124.86	120.30
1	B	198	ARG	CD-NE-CZ	9.11	136.35	123.60
1	B	183	ARG	CD-NE-CZ	8.53	135.54	123.60
1	A	183	ARG	CD-NE-CZ	8.47	135.45	123.60
1	B	286	ARG	NE-CZ-NH2	8.39	124.49	120.30
1	B	232	GLU	OE1-CD-OE2	-5.82	116.32	123.30
1	A	287[A]	LEU	CA-CB-CG	5.59	128.16	115.30
1	A	287[B]	LEU	CA-CB-CG	5.59	128.16	115.30
1	A	232	GLU	OE1-CD-OE2	-5.58	116.61	123.30
1	B	287[A]	LEU	CA-CB-CG	5.54	128.04	115.30
1	B	287[B]	LEU	CA-CB-CG	5.54	128.04	115.30
1	A	198	ARG	CG-CD-NE	-5.23	100.81	111.80

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	183	ARG	Sidechain
1	B	183	ARG	Sidechain
1	B	198	ARG	Sidechain

## 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	2485	0	2460	27	0
1	B	2507	0	2450	32	0
2	C	75	0	54	1	0
2	D	80	0	59	1	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	2	0	0	2	0
5	A	20	0	0	0	0
5	B	15	0	0	2	0
6	A	16	0	24	0	0
7	B	13	0	18	1	0
8	A	107	0	0	3	0
8	B	100	0	0	3	0
8	C	5	0	0	0	0
8	D	8	0	0	0	0
All	All	5435	0	5065	63	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 6.

All (63) 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:154:LEU:HD13	1:B:154:LEU:C	2.17	0.64
1:B:287[A]:LEU:HD23	1:B:288:SER:O	2.00	0.62
1:B:49:THR:HG21	1:B:114:LEU:HD23	1.82	0.62
1:B:51:LYS:NZ	8:B:367:HOH:O	2.33	0.61
1:A:154:LEU:C	1:A:154:LEU:HD13	2.21	0.60
1:A:323:GLY:HA3	1:A:329:TYR:CD2	2.41	0.56
1:A:154:LEU:O	1:A:157:ILE:HG22	2.05	0.56
1:B:248:ARG:NH2	8:B:368:HOH:O	2.39	0.55
1:B:323:GLY:HA3	1:B:329:TYR:CD2	2.42	0.55
1:A:160:ILE:HD12	1:A:190:THR:HG23	1.87	0.54
1:A:264:LEU:HD22	1:A:268:GLU:HB3	1.89	0.54
1:A:121:ALA:HB2	1:A:146:LEU:HD13	1.89	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:53:MET:HG2	1:B:78:LEU:HD22	1.91	0.52
1:A:297:ILE:O	1:A:308:GLN:HA	2.09	0.52
1:A:154:LEU:HD11	1:A:158:LYS:HD2	1.92	0.51
1:B:154:LEU:O	1:B:157:ILE:HG22	2.11	0.51
1:B:291:ARG:HD2	1:B:310:ILE:CD1	2.42	0.51
1:B:295:TRP:CE2	1:B:316:LEU:HD22	2.45	0.50
1:B:141:ARG:NE	5:B:5:SO4:O3	2.29	0.50
1:B:160:ILE:HD12	1:B:190:THR:HG23	1.94	0.50
1:B:124:LEU:HD21	1:B:142:ASN:HB2	1.94	0.50
1:B:295:TRP:CD2	1:B:316:LEU:HD22	2.46	0.49
1:B:70:SER:OG	2:D:1068:ARG:NH2	2.42	0.49
1:A:110:TYR:CE2	1:A:114:LEU:HD22	2.48	0.48
4:A:3:CL:CL	8:A:409:HOH:O	2.57	0.48
1:B:264:LEU:HD22	1:B:268:GLU:HB3	1.96	0.48
1:A:264:LEU:HD22	1:A:268:GLU:CB	2.44	0.47
1:B:271:ALA:O	1:B:274:GLN:HB2	2.14	0.47
1:B:269:VAL:HG21	1:B:286:ARG:HE	1.80	0.47
1:A:227:TYR:O	1:A:228:ILE:HD13	2.15	0.46
1:A:70:SER:OG	2:C:1068:ARG:NH2	2.38	0.46
1:B:154:LEU:HD11	1:B:158:LYS:HD2	1.97	0.46
1:A:287[A]:LEU:HD21	1:A:292:LEU:CD2	2.46	0.46
1:A:270:LYS:O	1:A:274:GLN:HG2	2.17	0.45
1:A:180:GLU:O	1:A:183:ARG:HB3	2.16	0.45
1:A:270:LYS:HA	1:A:306:ILE:HG21	1.99	0.45
5:B:3:SO4:O4	8:B:406:HOH:O	2.19	0.45
1:B:49:THR:HG21	1:B:114:LEU:CD2	2.46	0.45
1:B:161:PHE:CE2	1:B:166:PHE:HB2	2.52	0.45
1:A:287[A]:LEU:HD23	1:A:288:SER:O	2.17	0.44
1:A:160:ILE:HD12	1:A:190:THR:CG2	2.47	0.44
1:A:49:THR:HG21	1:A:114:LEU:HD23	1.98	0.44
1:A:53:MET:HG2	1:A:78:LEU:HD22	1.99	0.44
4:A:3:CL:CL	8:A:410:HOH:O	2.59	0.43
1:B:172[A]:ARG:HA	1:B:172[A]:ARG:HD3	1.72	0.43
1:A:248[A]:ARG:HD2	8:A:433:HOH:O	2.18	0.43
7:B:11:PG4:H82	7:B:11:PG4:H61	1.68	0.43
1:B:264:LEU:HD22	1:B:268:GLU:CB	2.49	0.43
1:B:124:LEU:HD11	1:B:142:ASN:HD22	1.84	0.43
1:B:121:ALA:HB2	1:B:146:LEU:HD13	2.01	0.43
1:B:110:TYR:CE2	1:B:114:LEU:HD22	2.54	0.42
1:A:271:ALA:O	1:A:274:GLN:HB2	2.19	0.42
1:B:49:THR:CG2	1:B:114:LEU:HG	2.50	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:172[B]:ARG:H	1:B:172[B]:ARG:HG2	1.53	0.42
1:A:62:ASN:ND2	1:A:133:TYR:OH	2.54	0.41
1:A:300:VAL:HG22	1:A:306:ILE:CD1	2.51	0.41
1:B:287[A]:LEU:HD21	1:B:292:LEU:CD2	2.50	0.41
1:B:84:HIS:O	1:B:88:ILE:HG13	2.20	0.41
1:A:261:MET:HB3	1:A:264:LEU:HD12	2.02	0.41
1:B:239:LEU:HD21	1:B:287[B]:LEU:HD13	2.02	0.40
1:A:124:LEU:HD21	1:A:139:ASP:O	2.21	0.40
1:A:291:ARG:HD2	1:A:310:ILE:CD1	2.52	0.40
1:B:65:LEU:HD23	1:B:140:ARG:HH11	1.85	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	304/315 (96%)	295 (97%)	9 (3%)	0	100	100
1	B	311/315 (99%)	295 (95%)	15 (5%)	1 (0%)	41	49
2	C	7/11 (64%)	6 (86%)	1 (14%)	0	100	100
2	D	8/11 (73%)	8 (100%)	0	0	100	100
All	All	630/652 (97%)	604 (96%)	25 (4%)	1 (0%)	47	57

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	346	GLU

### 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	266/281 (95%)	265 (100%)	1 (0%)	91	95
1	B	261/281 (93%)	260 (100%)	1 (0%)	91	95
2	C	6/8 (75%)	6 (100%)	0	100	100
2	D	6/8 (75%)	6 (100%)	0	100	100
All	All	539/578 (93%)	537 (100%)	2 (0%)	91	95

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	286	ARG
1	B	286	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	62	ASN
1	A	294	GLN
1	A	312	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	PTR	D	1069	2	15,16,17	2.20	1 (6%)	19,22,24	0.71	0
2	PTR	C	1069	2	15,16,17	2.00	2 (13%)	19,22,24	0.66	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	PTR	D	1069	2	-	0/10/11/13	0/1/1/1
2	PTR	C	1069	2	-	0/10/11/13	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1069	PTR	OH-CZ	-8.22	1.22	1.40
2	C	1069	PTR	OH-CZ	-6.86	1.25	1.40
2	C	1069	PTR	P-OH	2.68	1.63	1.59

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.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 4 are monoatomic - leaving 12 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The

Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	EDO	A	8[B]	-	3,3,3	0.50	0	2,2,2	0.35	0
6	EDO	A	8[A]	-	3,3,3	0.54	0	2,2,2	0.08	0
5	SO4	B	1	-	4,4,4	0.15	0	6,6,6	0.74	0
5	SO4	A	6	-	4,4,4	0.17	0	6,6,6	0.26	0
6	EDO	A	10	-	3,3,3	0.52	0	2,2,2	0.51	0
5	SO4	A	352	-	4,4,4	0.09	0	6,6,6	0.37	0
5	SO4	B	5	-	4,4,4	0.21	0	6,6,6	0.35	0
5	SO4	A	7	-	4,4,4	0.29	0	6,6,6	0.45	0
7	PG4	B	11	-	12,12,12	0.64	0	11,11,11	0.53	0
6	EDO	A	9	-	3,3,3	0.50	0	2,2,2	0.34	0
5	SO4	B	3	-	4,4,4	0.19	0	6,6,6	0.22	0
5	SO4	A	2	-	4,4,4	0.08	0	6,6,6	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
6	EDO	A	8[B]	-	-	0/1/1/1	-
6	EDO	A	9	-	-	0/1/1/1	-
6	EDO	A	8[A]	-	-	0/1/1/1	-
6	EDO	A	10	-	-	1/1/1/1	-
7	PG4	B	11	-	-	7/10/10/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	11	PG4	C8-C7-O4-C6
7	B	11	PG4	O2-C3-C4-O3
7	B	11	PG4	O1-C1-C2-O2
7	B	11	PG4	C4-C3-O2-C2
7	B	11	PG4	C6-C5-O3-C4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
6	A	10	EDO	O1-C1-C2-O2
7	B	11	PG4	C5-C6-O4-C7
7	B	11	PG4	O3-C5-C6-O4

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	5	SO4	1	0
7	B	11	PG4	1	0
5	B	3	SO4	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	302/315 (95%)	0.40	16 (5%)	26 31	24, 41, 87, 111	0
1	B	310/315 (98%)	0.55	24 (7%)	13 16	23, 43, 92, 119	0
2	C	9/11 (81%)	1.11	2 (22%)	0 1	41, 43, 78, 88	0
2	D	10/11 (90%)	0.60	2 (20%)	1 1	46, 66, 96, 98	0
All	All	631/652 (96%)	0.49	44 (6%)	16 20	23, 43, 90, 119	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	95	ASN	8.4
1	B	96	GLN	7.8
1	B	133	TYR	7.0
1	B	99	ALA	6.8
1	B	94	ASP	5.3
1	B	349	TYR	4.4
2	C	1066	LEU	4.2
2	C	1067	GLN	4.0
1	A	61	GLN	3.7
1	B	312	HIS	3.6
1	A	133	TYR	3.6
1	B	183	ARG	3.3
1	A	96	GLN	3.0
1	A	62	ASN	2.9
1	B	98	LEU	2.9
1	B	186	PHE	2.8
1	A	95	ASN	2.8
1	B	100	GLN	2.7
1	B	55	LYS	2.6
1	A	179	ALA	2.5
1	B	132	MET	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	127	GLU	2.5
1	B	93	ASP	2.4
1	A	278	THR	2.4
1	B	131	ARG	2.4
1	A	105	GLU	2.4
1	B	188	ASP	2.4
2	D	1076	ALA	2.3
1	B	102	SER	2.3
1	B	92	TYR	2.3
1	A	45	THR	2.3
1	A	161	PHE	2.3
1	A	336	SER	2.3
1	B	277	SER	2.3
1	B	160	ILE	2.2
2	D	1066	LEU	2.2
1	A	63	PRO	2.2
1	B	50	TRP	2.2
1	B	129	LYS	2.2
1	A	94	ASP	2.2
1	B	82	TYR	2.1
1	A	176	ALA	2.1
1	A	66	GLN	2.1
1	A	64	LYS	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
2	PTR	D	1069	16/17	0.92	0.18	47,51,62,63	0
2	PTR	C	1069	16/17	0.97	0.14	33,36,40,43	0

## 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. 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(Å <sup>2</sup> )	Q<0.9
7	PG4	B	11	13/13	0.76	0.30	37,49,63,65	0
6	EDO	A	8[A]	4/4	0.78	0.21	47,47,49,54	4
6	EDO	A	8[B]	4/4	0.78	0.21	48,51,54,60	4
6	EDO	A	10	4/4	0.83	0.26	57,58,61,62	0
6	EDO	A	9	4/4	0.86	0.40	64,66,67,70	0
5	SO4	B	5	5/5	0.91	0.12	67,68,75,77	0
4	CL	A	3	1/1	0.91	0.05	67,67,67,67	0
5	SO4	A	6	5/5	0.92	0.11	68,69,77,77	0
3	NA	B	2	1/1	0.93	0.10	41,41,41,41	0
5	SO4	A	7	5/5	0.93	0.17	54,55,68,74	0
5	SO4	B	1	5/5	0.94	0.14	38,54,62,62	0
3	NA	A	1	1/1	0.94	0.18	35,35,35,35	0
4	CL	A	4	1/1	0.95	0.17	60,60,60,60	0
5	SO4	A	352	5/5	0.97	0.11	47,57,61,62	0
5	SO4	B	3	5/5	0.97	0.20	46,47,58,63	0
5	SO4	A	2	5/5	0.97	0.10	42,51,74,80	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.