



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:32 PM GMT

PDB ID : 3ZNI  
Title : Structure of phosphoTyr363-Cbl-b - UbcH5B-Ub - ZAP-70 peptide complex  
Authors : Dou, H.; Buetow, L.; Sibbet, G.J.; Cameron, K.; Huang, D.T.  
Deposited on : 2013-02-14  
Resolution : 2.21 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

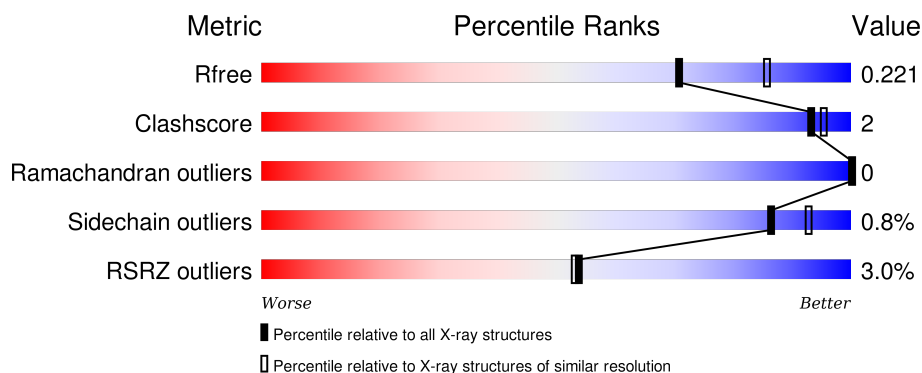
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.21 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	4405 (2.24-2.20)
Clashscore	102246	5146 (2.24-2.20)
Ramachandran outliers	100387	5065 (2.24-2.20)
Sidechain outliers	100360	5066 (2.24-2.20)
RSRZ outliers	91569	4414 (2.24-2.20)

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	394	<div> <div>92%</div> <div>7%</div> </div>
1	E	394	<div> <div>3%</div> <div>94%</div> <div>5%</div> </div>
1	I	394	<div> <div>4%</div> <div>91%</div> <div>7%</div> </div>
1	M	394	<div> <div>3%</div> <div>94%</div> <div>5%</div> </div>
2	B	12	<div> <div>25%</div> <div>75%</div> <div>25%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	12	
2	J	12	
2	N	12	
3	C	146	
3	G	146	
3	K	146	
3	O	146	
4	D	81	
4	H	81	
4	L	81	
4	P	81	

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
7	EDO	E	1431	-	-	-	X
7	EDO	M	1431	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 20546 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	390	Total	C	N	O	P	S	0	2	0
			3098	1987	522	565	1	23			
1	E	390	Total	C	N	O	P	S	0	1	0
			3120	2002	526	568	1	23			
1	I	386	Total	C	N	O	P	S	0	1	0
			3072	1969	517	563	1	22			
1	M	390	Total	C	N	O	P	S	0	3	0
			3128	2006	531	568	1	22			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	GLY	-	EXPRESSION TAG	UNP Q13191
A	35	SER	-	EXPRESSION TAG	UNP Q13191
A	360	PHE	TYR	ENGINEERED MUTATION	UNP Q13191
E	34	GLY	-	EXPRESSION TAG	UNP Q13191
E	35	SER	-	EXPRESSION TAG	UNP Q13191
E	360	PHE	TYR	ENGINEERED MUTATION	UNP Q13191
I	34	GLY	-	EXPRESSION TAG	UNP Q13191
I	35	SER	-	EXPRESSION TAG	UNP Q13191
I	360	PHE	TYR	ENGINEERED MUTATION	UNP Q13191
M	34	GLY	-	EXPRESSION TAG	UNP Q13191
M	35	SER	-	EXPRESSION TAG	UNP Q13191
M	360	PHE	TYR	ENGINEERED MUTATION	UNP Q13191

- Molecule 2 is a protein called TYROSINE-PROTEIN KINASE ZAP-70.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	9	Total	C	N	O	P	0	0	0
			69	40	9	19	1			
2	F	8	Total	C	N	O	P	0	0	0
			63	37	8	17	1			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	8	Total	C	N	O	P	0	0	0
			63	37	8	17	1			
2	N	9	Total	C	N	O	P	0	0	0
			68	40	9	18	1			

- Molecule 3 is a protein called UBIQUITIN-CONJUGATING ENZYME E2 D2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	146	Total	C	N	O	S	0	0	0
			1162	748	198	210	6			
3	G	146	Total	C	N	O	S	0	0	0
			1172	753	202	211	6			
3	K	146	Total	C	N	O	S	0	1	0
			1179	756	205	212	6			
3	O	146	Total	C	N	O	S	0	3	0
			1187	762	207	212	6			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	22	ARG	SER	ENGINEERED MUTATION	UNP P62837
C	85	LYS	CYS	ENGINEERED MUTATION	UNP P62837
G	22	ARG	SER	ENGINEERED MUTATION	UNP P62837
G	85	LYS	CYS	ENGINEERED MUTATION	UNP P62837
K	22	ARG	SER	ENGINEERED MUTATION	UNP P62837
K	85	LYS	CYS	ENGINEERED MUTATION	UNP P62837
O	22	ARG	SER	ENGINEERED MUTATION	UNP P62837
O	85	LYS	CYS	ENGINEERED MUTATION	UNP P62837

- Molecule 4 is a protein called POLYUBIQUITIN-C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	77	Total	C	N	O	S	0	0	0
			599	375	104	119	1			
4	H	77	Total	C	N	O	S	0	0	0
			595	374	105	115	1			
4	L	77	Total	C	N	O	S	0	0	0
			583	368	101	113	1			
4	P	77	Total	C	N	O	S	0	0	0
			600	377	104	118	1			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-4	GLY	-	EXPRESSION TAG	UNP P0CG48
D	-3	SER	-	EXPRESSION TAG	UNP P0CG48
D	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
D	-1	GLY	-	EXPRESSION TAG	UNP P0CG48
D	0	SER	-	EXPRESSION TAG	UNP P0CG48
H	-4	GLY	-	EXPRESSION TAG	UNP P0CG48
H	-3	SER	-	EXPRESSION TAG	UNP P0CG48
H	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
H	-1	GLY	-	EXPRESSION TAG	UNP P0CG48
H	0	SER	-	EXPRESSION TAG	UNP P0CG48
L	-4	GLY	-	EXPRESSION TAG	UNP P0CG48
L	-3	SER	-	EXPRESSION TAG	UNP P0CG48
L	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
L	-1	GLY	-	EXPRESSION TAG	UNP P0CG48
L	0	SER	-	EXPRESSION TAG	UNP P0CG48
P	-4	GLY	-	EXPRESSION TAG	UNP P0CG48
P	-3	SER	-	EXPRESSION TAG	UNP P0CG48
P	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
P	-1	GLY	-	EXPRESSION TAG	UNP P0CG48
P	0	SER	-	EXPRESSION TAG	UNP P0CG48

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	I	2	Total Zn 2 2	0	0
5	A	2	Total Zn 2 2	0	0
5	M	2	Total Zn 2 2	0	0
5	E	2	Total Zn 2 2	0	0

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	I	1	Total Ca 1 1	0	0
6	A	1	Total Ca 1 1	0	0
6	M	1	Total Ca 1 1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	E	1	Total	Ca	0	0
			1	1		

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			4	2	2		
7	E	1	Total	C	O	0	0
			4	2	2		
7	I	1	Total	C	O	0	0
			4	2	2		
7	M	1	Total	C	O	0	0
			4	2	2		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	125	Total	O	0	0
			125	125		
8	B	1	Total	O	0	0
			1	1		
8	C	67	Total	O	0	0
			67	67		
8	D	14	Total	O	0	0
			14	14		

*Continued on next page...*

*Continued from previous page...*

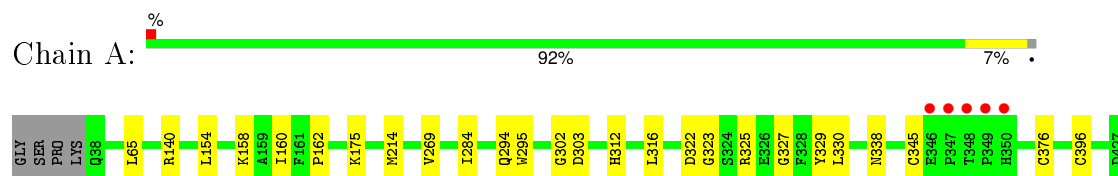
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	E	171	Total 171	O 171	0	0
8	F	1	Total 1	O 1	0	0
8	G	49	Total 49	O 49	0	0
8	H	7	Total 7	O 7	0	0
8	I	98	Total 98	O 98	0	0
8	J	2	Total 2	O 2	0	0
8	K	37	Total 37	O 37	0	0
8	L	7	Total 7	O 7	0	0
8	M	111	Total 111	O 111	0	0
8	O	60	Total 60	O 60	0	0
8	P	10	Total 10	O 10	0	0



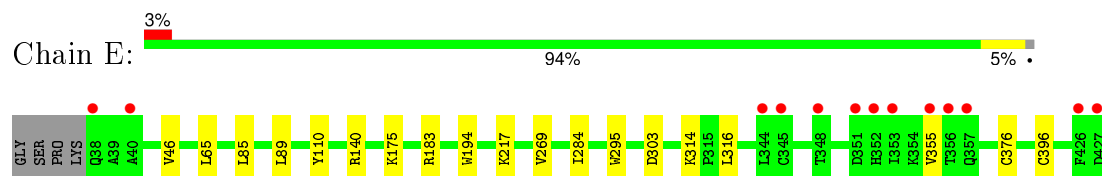
### 3 Residue-property plots [i](#)

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

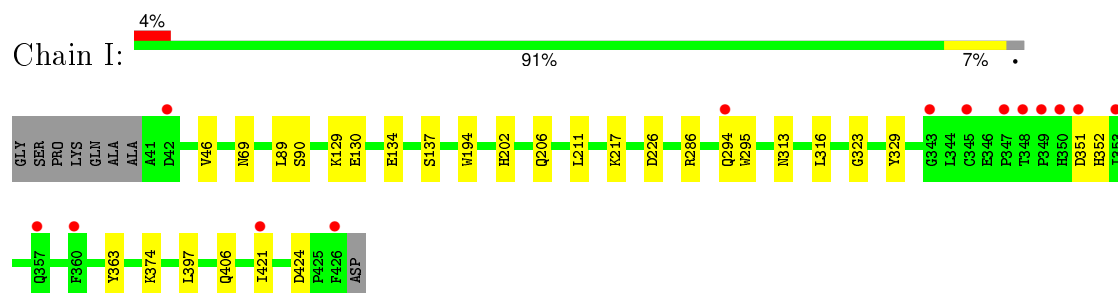
- Molecule 1: E3 UBIQUITIN-PROTEIN LIGASE CBL-B



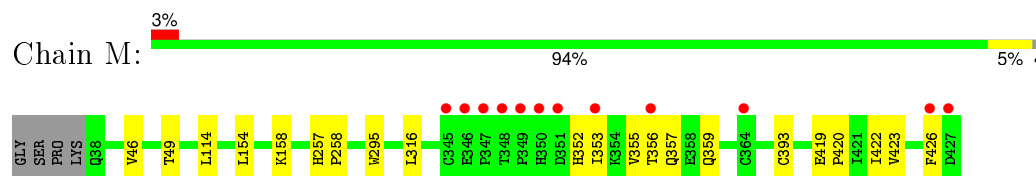
- Molecule 1: E3 UBIQUITIN-PROTEIN LIGASE CBL-B



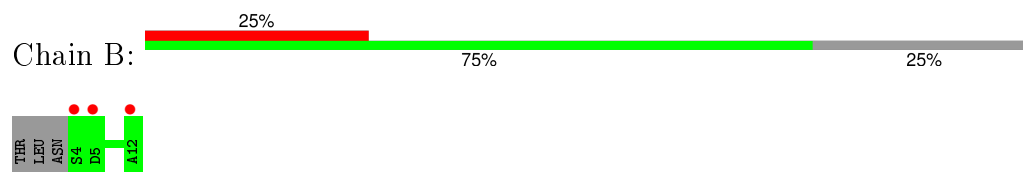
- Molecule 1: E3 UBIQUITIN-PROTEIN LIGASE CBL-B



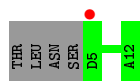
- Molecule 1: E3 UBIQUITIN-PROTEIN LIGASE CBL-B



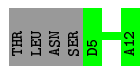
- Molecule 2: TYROSINE-PROTEIN KINASE ZAP-70



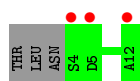
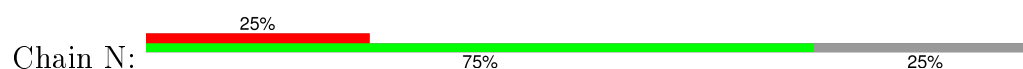
- Molecule 2: TYROSINE-PROTEIN KINASE ZAP-70



- Molecule 2: TYROSINE-PROTEIN KINASE ZAP-70



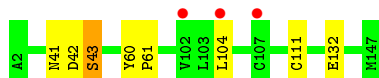
- Molecule 2: TYROSINE-PROTEIN KINASE ZAP-70



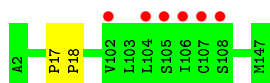
- Molecule 3: UBIQUITIN-CONJUGATING ENZYME E2 D2



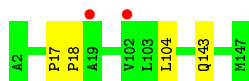
- Molecule 3: UBIQUITIN-CONJUGATING ENZYME E2 D2




- Molecule 3: UBIQUITIN-CONJUGATING ENZYME E2 D2



- Molecule 3: UBIQUITIN-CONJUGATING ENZYME E2 D2




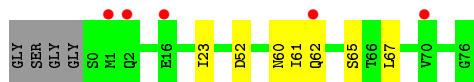
- Molecule 4: POLYUBIQUITIN-C

Chain D:  86% 9% 5%




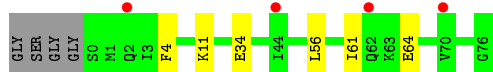
- Molecule 4: POLYUBIQUITIN-C

Chain H:  86% 9% 5% 6%




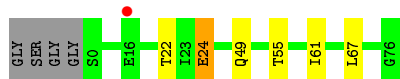
- Molecule 4: POLYUBIQUITIN-C

Chain L:  88% 7% 5% 5%



- Molecule 4: POLYUBIQUITIN-C

Chain P:  88% 6% 5% 0%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.97Å 131.80Å 122.00Å 90.00° 91.92° 90.00°	Depositor
Resolution (Å)	30.54 – 2.21 30.54 – 2.21	Depositor EDS
% Data completeness (in resolution range)	99.6 (30.54-2.21) 99.6 (30.54-2.21)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.33 (at 2.22Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.175 , 0.211 0.187 , 0.221	Depositor DCC
$R_{free}$ test set	7485 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	42.2	Xtriage
Anisotropy	0.365	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 34.7	EDS
Estimated twinning fraction	0.022 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 149245 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	20546	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.68% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, CA, PTR, EDO

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	0.27	0/3164	0.41	0/4291
1	E	0.29	0/3183	0.43	0/4310
1	I	0.26	0/3136	0.41	0/4252
1	M	0.27	0/3198	0.41	0/4331
2	B	0.22	0/53	0.43	0/71
2	F	0.24	0/47	0.38	0/63
2	J	0.21	0/47	0.39	0/63
2	N	0.24	0/52	0.40	0/70
3	C	0.27	0/1198	0.44	0/1634
3	G	0.26	0/1208	0.45	0/1646
3	K	0.25	0/1218	0.44	0/1660
3	O	0.28	0/1232	0.48	0/1678
4	D	0.24	0/605	0.46	0/816
4	H	0.24	0/601	0.46	0/810
4	L	0.25	0/589	0.44	0/796
4	P	0.26	0/606	0.44	0/816
All	All	0.27	0/20137	0.43	0/27307

There are no bond length outliers.

There are no bond angle outliers.

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	3098	0	2967	14	0
1	E	3120	0	3022	8	0
1	I	3072	0	2937	16	0
1	M	3128	0	3024	13	0
2	B	69	0	50	0	0
2	F	63	0	45	0	0
2	J	63	0	45	0	0
2	N	68	0	47	0	0
3	C	1162	0	1139	3	0
3	G	1172	0	1156	5	0
3	K	1179	0	1162	1	0
3	O	1187	0	1176	2	0
4	D	599	0	612	3	0
4	H	595	0	615	3	0
4	L	583	0	593	5	0
4	P	600	0	619	3	0
5	A	2	0	0	0	0
5	E	2	0	0	0	0
5	I	2	0	0	0	0
5	M	2	0	0	0	0
6	A	1	0	0	0	0
6	E	1	0	0	0	0
6	I	1	0	0	0	0
6	M	1	0	0	0	0
7	A	4	0	6	0	0
7	E	4	0	6	0	0
7	I	4	0	6	0	0
7	M	4	0	6	0	0
8	A	125	0	0	1	0
8	B	1	0	0	0	0
8	C	67	0	0	0	0
8	D	14	0	0	0	0
8	E	171	0	0	0	0
8	F	1	0	0	0	0
8	G	49	0	0	0	0
8	H	7	0	0	0	0
8	I	98	0	0	0	0
8	J	2	0	0	0	0
8	K	37	0	0	0	0
8	L	7	0	0	0	0
8	M	111	0	0	0	0
8	O	60	0	0	0	0
8	P	10	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	20546	0	19233	74	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.

The worst 5 of 74 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:A:65:LEU:O	1:A:140:ARG:NH2	2.16	0.78
1:E:46:VAL:HG11	1:E:89:LEU:HD11	1.68	0.75
3:G:132:GLU:HG2	1:I:90:SER:HA	1.69	0.73
4:D:14:THR:O	4:D:33:LYS:NZ	2.23	0.71
1:I:46:VAL:HG11	1:I:89:LEU:HD11	1.76	0.66

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

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

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	389/394 (99%)	380 (98%)	9 (2%)	0	100	100
1	E	388/394 (98%)	381 (98%)	7 (2%)	0	100	100
1	I	384/394 (98%)	376 (98%)	8 (2%)	0	100	100
1	M	390/394 (99%)	381 (98%)	9 (2%)	0	100	100
2	B	6/12 (50%)	6 (100%)	0	0	100	100
2	F	5/12 (42%)	5 (100%)	0	0	100	100
2	J	5/12 (42%)	5 (100%)	0	0	100	100
2	N	6/12 (50%)	6 (100%)	0	0	100	100
3	C	144/146 (99%)	142 (99%)	2 (1%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	G	144/146 (99%)	140 (97%)	4 (3%)	0	100	100
3	K	145/146 (99%)	143 (99%)	2 (1%)	0	100	100
3	O	147/146 (101%)	143 (97%)	4 (3%)	0	100	100
4	D	75/81 (93%)	73 (97%)	2 (3%)	0	100	100
4	H	75/81 (93%)	74 (99%)	1 (1%)	0	100	100
4	L	75/81 (93%)	73 (97%)	2 (3%)	0	100	100
4	P	75/81 (93%)	74 (99%)	1 (1%)	0	100	100
All	All	2453/2532 (97%)	2402 (98%)	51 (2%)	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	325/352 (92%)	324 (100%)	1 (0%)	94	98
1	E	332/352 (94%)	328 (99%)	4 (1%)	78	88
1	I	325/352 (92%)	321 (99%)	4 (1%)	78	88
1	M	332/352 (94%)	331 (100%)	1 (0%)	94	98
2	B	6/9 (67%)	6 (100%)	0	100	100
2	F	5/9 (56%)	5 (100%)	0	100	100
2	J	5/9 (56%)	5 (100%)	0	100	100
2	N	5/9 (56%)	5 (100%)	0	100	100
3	C	127/130 (98%)	127 (100%)	0	100	100
3	G	129/130 (99%)	128 (99%)	1 (1%)	86	93
3	K	130/130 (100%)	130 (100%)	0	100	100
3	O	131/130 (101%)	130 (99%)	1 (1%)	86	93
4	D	67/70 (96%)	66 (98%)	1 (2%)	72	84
4	H	66/70 (94%)	65 (98%)	1 (2%)	72	84

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	L	63/70 (90%)	63 (100%)	0	100	100
4	P	67/70 (96%)	65 (97%)	2 (3%)	48	60
All	All	2115/2244 (94%)	2099 (99%)	16 (1%)	86	93

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	H	60	ASN
1	I	211	LEU
1	M	393	CYS
3	G	43	SER
3	O	143	GLN

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	294	GLN
4	H	62	GLN
1	M	352	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 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	$\# Z  > 2$
1	PTR	A	363	1	14,16,17	1.21	1 (7%)	18,22,24	0.73	1 (5%)
2	PTR	B	7	2	14,16,17	1.17	1 (7%)	18,22,24	0.65	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	PTR	E	363	1	14,16,17	1.26	1 (7%)	18,22,24	0.66	1 (5%)
2	PTR	F	7	2	14,16,17	1.24	1 (7%)	18,22,24	0.75	0
1	PTR	I	363	1	14,16,17	1.19	1 (7%)	18,22,24	0.64	0
2	PTR	J	7	2	14,16,17	1.23	1 (7%)	18,22,24	0.68	0
1	PTR	M	363	1	14,16,17	1.22	1 (7%)	18,22,24	0.61	0
2	PTR	N	7	2	14,16,17	1.26	1 (7%)	18,22,24	0.63	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
1	PTR	A	363	1	-	0/9/11/13	0/1/1/1
2	PTR	B	7	2	-	0/9/11/13	0/1/1/1
1	PTR	E	363	1	-	0/9/11/13	0/1/1/1
2	PTR	F	7	2	-	0/9/11/13	0/1/1/1
1	PTR	I	363	1	-	0/9/11/13	0/1/1/1
2	PTR	J	7	2	-	0/9/11/13	0/1/1/1
1	PTR	M	363	1	-	0/9/11/13	0/1/1/1
2	PTR	N	7	2	-	0/9/11/13	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	N	7	PTR	OH-CZ	-4.10	1.30	1.40
1	E	363	PTR	OH-CZ	-4.09	1.30	1.40
2	F	7	PTR	OH-CZ	-4.07	1.30	1.40
2	J	7	PTR	OH-CZ	-4.04	1.30	1.40
1	A	363	PTR	OH-CZ	-4.00	1.31	1.40

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	363	PTR	O-C-CA	-2.16	119.87	125.49
1	A	363	PTR	O-C-CA	-2.10	120.01	125.49

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	I	363	PTR	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 12 are monoatomic - leaving 4 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	$\# Z  > 2$
7	EDO	A	1431	-	3,3,3	0.48	0	2,2,2	0.39	0
7	EDO	E	1431	-	3,3,3	0.56	0	2,2,2	0.21	0
7	EDO	I	1430	-	3,3,3	0.47	0	2,2,2	0.39	0
7	EDO	M	1431	-	3,3,3	0.46	0	2,2,2	0.39	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
7	EDO	A	1431	-	-	0/1/1/1	0/0/0/0
7	EDO	E	1431	-	-	0/1/1/1	0/0/0/0
7	EDO	I	1430	-	-	0/1/1/1	0/0/0/0
7	EDO	M	1431	-	-	0/1/1/1	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	389/394 (98%)	-0.29	5 (1%) 79 78	24, 43, 80, 136	0
1	E	389/394 (98%)	-0.20	13 (3%) 50 49	24, 40, 78, 113	0
1	I	385/394 (97%)	-0.09	14 (3%) 46 45	26, 48, 85, 141	0
1	M	389/394 (98%)	-0.21	12 (3%) 52 51	30, 45, 85, 135	0
2	B	8/12 (66%)	1.37	3 (37%) 0 0	59, 67, 82, 94	0
2	F	7/12 (58%)	0.12	1 (14%) 4 3	42, 48, 63, 70	0
2	J	7/12 (58%)	0.29	0 100 100	52, 58, 69, 75	0
2	N	8/12 (66%)	1.86	3 (37%) 0 0	54, 64, 92, 105	0
3	C	146/146 (100%)	-0.32	2 (1%) 78 77	23, 35, 59, 77	0
3	G	146/146 (100%)	-0.31	3 (2%) 67 65	28, 43, 71, 83	0
3	K	146/146 (100%)	-0.24	6 (4%) 41 39	30, 44, 71, 89	0
3	O	146/146 (100%)	-0.30	2 (1%) 78 77	22, 34, 60, 84	0
4	D	77/81 (95%)	-0.11	0 100 100	24, 49, 77, 81	0
4	H	77/81 (95%)	0.21	5 (6%) 22 21	27, 53, 87, 102	0
4	L	77/81 (95%)	0.30	4 (5%) 31 30	33, 60, 88, 100	0
4	P	77/81 (95%)	-0.24	1 (1%) 79 78	29, 49, 72, 81	0
All	All	2474/2532 (97%)	-0.18	74 (2%) 54 53	22, 44, 81, 141	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	345	CYS	8.6
1	I	348	THR	6.1
1	M	348	THR	5.9
2	N	4	SER	5.6
1	I	347	PRO	5.5

## 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. 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(Å <sup>2</sup> )	Q<0.9
1	PTR	E	363	16/17	0.97	0.10	-	40,48,59,60	0
2	PTR	J	7	16/17	0.98	0.13	-	35,40,49,51	0
1	PTR	I	363	16/17	0.87	0.15	-	57,62,67,67	0
1	PTR	M	363	16/17	0.96	0.10	-	47,61,69,71	0
2	PTR	B	7	16/17	0.97	0.11	-	41,46,59,60	0
2	PTR	F	7	16/17	0.97	0.11	-	31,34,39,40	0
2	PTR	N	7	16/17	0.97	0.09	-	36,41,51,51	0
1	PTR	A	363	16/17	0.94	0.12	-	44,52,64,65	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. 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(Å <sup>2</sup> )	Q<0.9
7	EDO	E	1431	4/4	0.83	0.28	3.07	43,44,46,48	0
7	EDO	M	1431	4/4	0.91	0.20	2.09	49,51,56,59	0
7	EDO	A	1431	4/4	0.96	0.17	1.66	41,41,41,46	0
5	ZN	E	1429	1/1	1.00	0.11	0.57	31,31,31,31	0
5	ZN	M	1429	1/1	1.00	0.12	0.04	37,37,37,37	0
7	EDO	I	1430	4/4	0.97	0.13	0.01	42,44,47,50	0
5	ZN	I	1428	1/1	1.00	0.10	-0.16	37,37,37,37	0
5	ZN	M	1428	1/1	0.99	0.08	-0.62	42,42,42,42	0
5	ZN	A	1428	1/1	0.99	0.08	-0.87	40,40,40,40	0
6	CA	E	1430	1/1	0.97	0.08	-0.91	52,52,52,52	0
5	ZN	E	1428	1/1	0.99	0.07	-0.95	38,38,38,38	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	ZN	A	1429	1/1	1.00	0.11	-1.27	28,28,28,28	0
6	CA	M	1430	1/1	0.97	0.08	-1.34	55,55,55,55	0
5	ZN	I	1427	1/1	0.99	0.06	-1.41	51,51,51,51	0
6	CA	A	1430	1/1	0.97	0.07	-1.65	51,51,51,51	0
6	CA	I	1429	1/1	0.97	0.04	-3.59	61,61,61,61	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.