



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 15, 2017 – 12:49 am GMT

PDB ID : 3I3J  
Title : Crystal Structure of the Bromodomain of Human EP300  
Authors : Filippakopoulos, P.; Picaud, S.; Phillips, C.; Pike, A.C.W.; Muniz, J.; Roos, A.; Chaikuad, A.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.M.; Weigelt, J.; Bountra, C.; Knapp, S.  
Deposited on : 2009-06-30  
Resolution : 2.33 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 : trunk28620  
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) : recalc28949

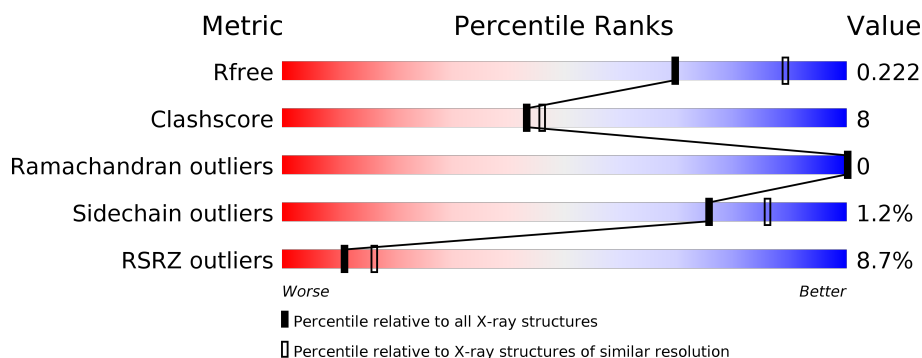
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.33 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	1570 (2.36-2.32)
Clashscore	112137	1673 (2.36-2.32)
Ramachandran outliers	110173	1654 (2.36-2.32)
Sidechain outliers	110143	1655 (2.36-2.32)
RSRZ outliers	101464	1576 (2.36-2.32)

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	124	<div> <div>6%</div> <div> <div></div> <div>70%</div> <div>22%</div> <div>8%</div> </div> </div>
1	B	124	<div> <div>6%</div> <div> <div></div> <div>80%</div> <div>11%</div> <div>8%</div> </div> </div>
1	C	124	<div> <div>6%</div> <div> <div></div> <div>73%</div> <div>17%</div> <div>8%</div> </div> </div>
1	D	124	<div> <div>6%</div> <div> <div></div> <div>73%</div> <div>16%</div> <div>9%</div> </div> </div>
1	E	124	<div> <div>3%</div> <div> <div></div> <div>74%</div> <div>19%</div> <div>7%</div> </div> </div>
1	F	124	<div> <div>24%</div> <div> <div></div> <div>75%</div> <div>14%</div> <div>10%</div> </div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	124	
1	H	124	
1	I	124	
1	J	124	
1	K	124	
1	L	124	

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	CL	A	3	-	-	-	X
3	PEG	A	1	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11399 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 Histone acetyltransferase p300.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	114	Total	C	N	O	S	0	0	0
			932	609	145	173	5			
1	B	114	Total	C	N	O	S	0	0	0
			913	598	139	171	5			
1	C	114	Total	C	N	O	S	0	0	0
			944	612	152	175	5			
1	D	113	Total	C	N	O	S	0	0	0
			920	604	139	172	5			
1	E	115	Total	C	N	O	S	0	0	0
			948	616	149	178	5			
1	F	111	Total	C	N	O	S	0	0	0
			890	580	139	166	5			
1	G	113	Total	C	N	O	S	0	0	0
			918	599	143	171	5			
1	H	113	Total	C	N	O	S	0	1	0
			922	602	145	170	5			
1	I	115	Total	C	N	O	S	0	0	0
			944	615	148	176	5			
1	J	113	Total	C	N	O	S	0	0	0
			925	605	142	173	5			
1	K	114	Total	C	N	O	S	0	1	0
			934	609	151	169	5			
1	L	113	Total	C	N	O	S	0	0	0
			906	590	146	165	5			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1038	SER	-	EXPRESSION TAG	UNP Q09472
A	1039	MET	-	EXPRESSION TAG	UNP Q09472
B	1038	SER	-	EXPRESSION TAG	UNP Q09472
B	1039	MET	-	EXPRESSION TAG	UNP Q09472
C	1038	SER	-	EXPRESSION TAG	UNP Q09472

*Continued on next page...*

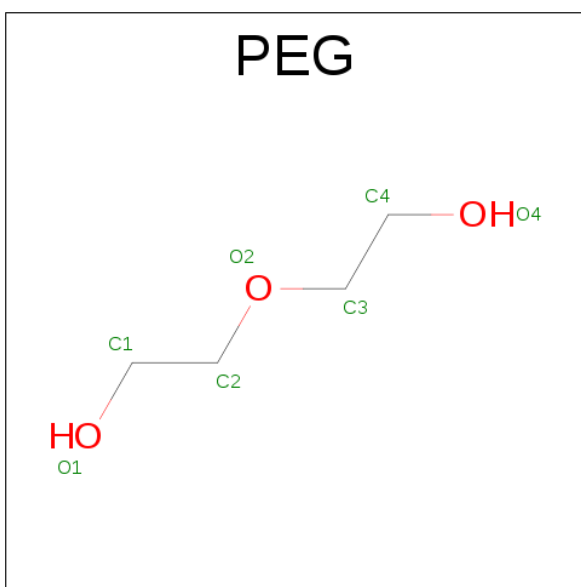
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	1039	MET	-	EXPRESSION TAG	UNP Q09472
D	1038	SER	-	EXPRESSION TAG	UNP Q09472
D	1039	MET	-	EXPRESSION TAG	UNP Q09472
E	1038	SER	-	EXPRESSION TAG	UNP Q09472
E	1039	MET	-	EXPRESSION TAG	UNP Q09472
F	1038	SER	-	EXPRESSION TAG	UNP Q09472
F	1039	MET	-	EXPRESSION TAG	UNP Q09472
G	1038	SER	-	EXPRESSION TAG	UNP Q09472
G	1039	MET	-	EXPRESSION TAG	UNP Q09472
H	1038	SER	-	EXPRESSION TAG	UNP Q09472
H	1039	MET	-	EXPRESSION TAG	UNP Q09472
I	1038	SER	-	EXPRESSION TAG	UNP Q09472
I	1039	MET	-	EXPRESSION TAG	UNP Q09472
J	1038	SER	-	EXPRESSION TAG	UNP Q09472
J	1039	MET	-	EXPRESSION TAG	UNP Q09472
K	1038	SER	-	EXPRESSION TAG	UNP Q09472
K	1039	MET	-	EXPRESSION TAG	UNP Q09472
L	1038	SER	-	EXPRESSION TAG	UNP Q09472
L	1039	MET	-	EXPRESSION TAG	UNP Q09472

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

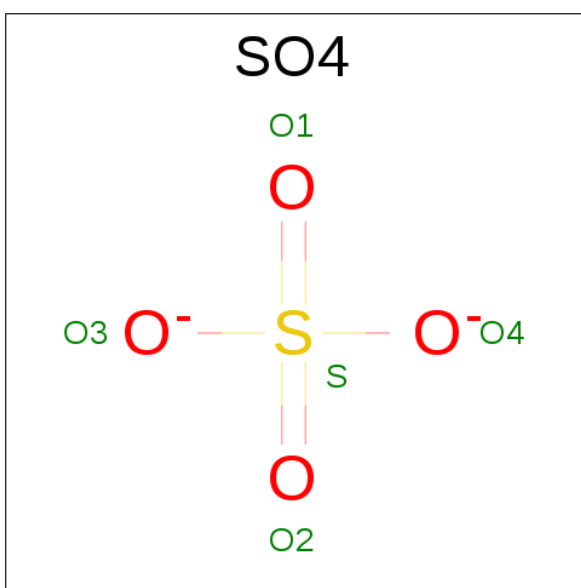
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Cl 1 1	0	0
2	E	1	Total Cl 1 1	0	0

- Molecule 3 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
3	A	1	Total	C	O	0	0
			7	4	3		
3	E	1	Total	C	O	0	0
			7	4	3		

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



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

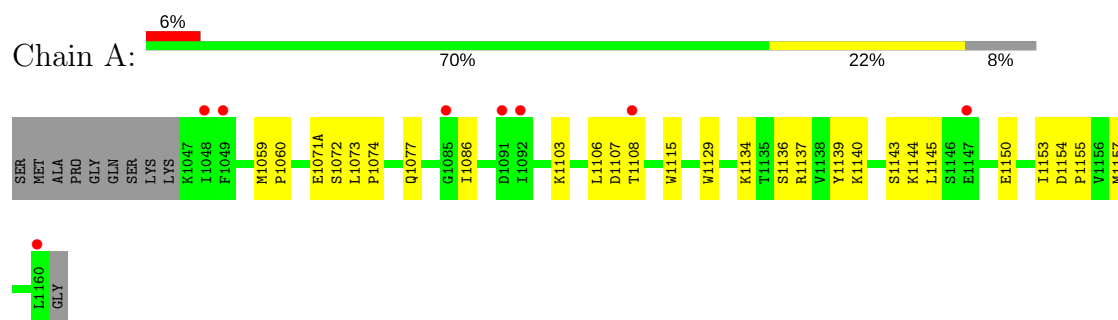
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	39	Total 39	O 39	0	0
5	B	14	Total 14	O 14	0	0
5	C	31	Total 31	O 31	0	0
5	D	20	Total 20	O 20	0	0
5	E	32	Total 32	O 32	0	0
5	F	1	Total 1	O 1	0	0
5	G	38	Total 38	O 38	0	0
5	H	15	Total 15	O 15	0	0
5	I	24	Total 24	O 24	0	0
5	J	23	Total 23	O 23	0	0
5	K	28	Total 28	O 28	0	0
5	L	12	Total 12	O 12	0	0

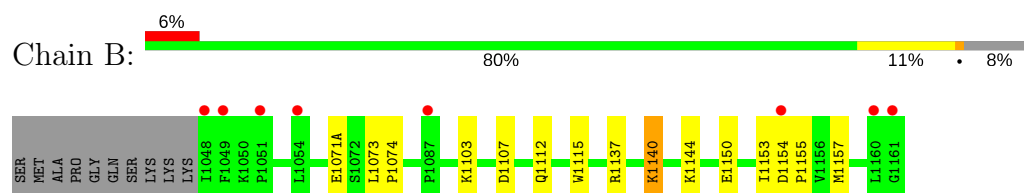
### 3 Residue-property plots [i](#)

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

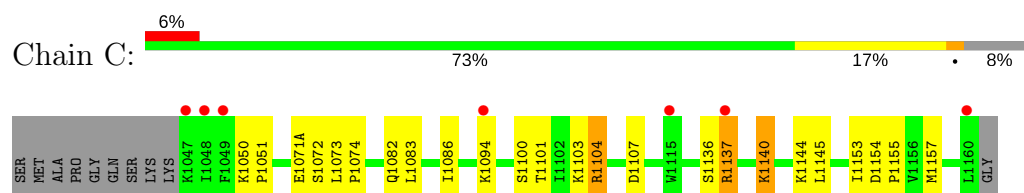
#### • Molecule 1: Histone acetyltransferase p300



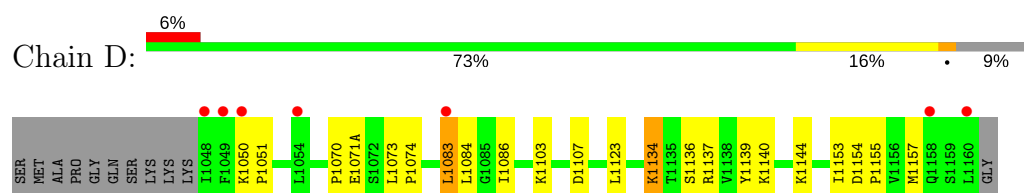
#### • Molecule 1: Histone acetyltransferase p300



#### • Molecule 1: Histone acetyltransferase p300



#### • Molecule 1: Histone acetyltransferase p300

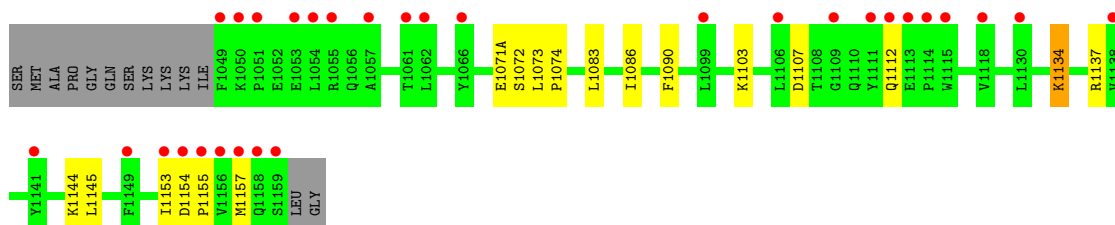
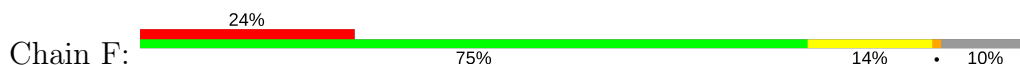


#### • Molecule 1: Histone acetyltransferase p300

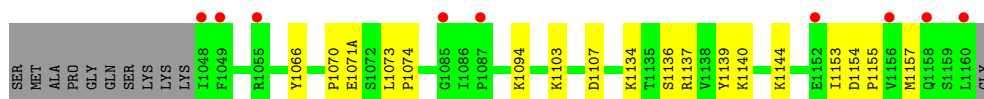
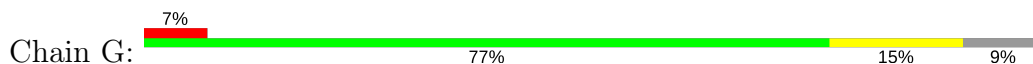




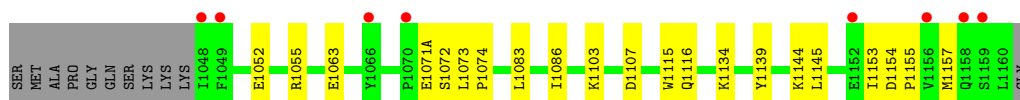
- Molecule 1: Histone acetyltransferase p300



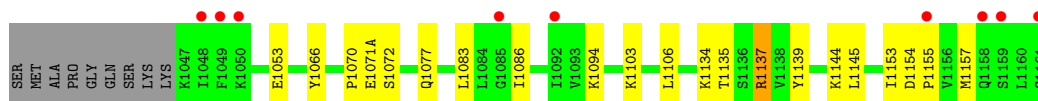
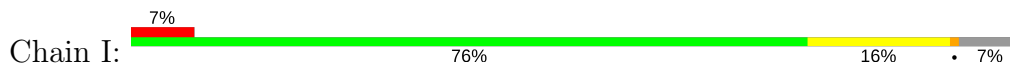
- Molecule 1: Histone acetyltransferase p300



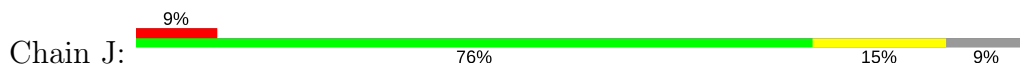
- Molecule 1: Histone acetyltransferase p300



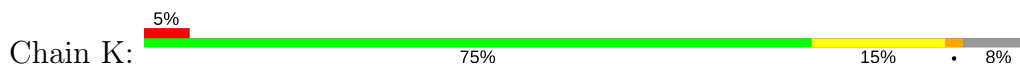
- Molecule 1: Histone acetyltransferase p300



- Molecule 1: Histone acetyltransferase p300

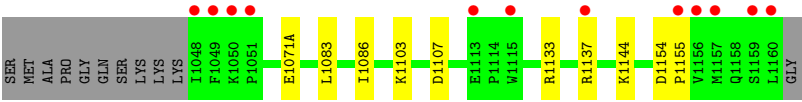
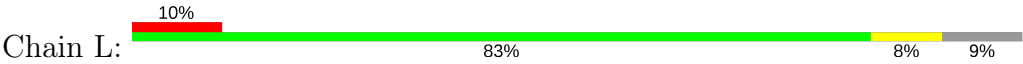


- Molecule 1: Histone acetyltransferase p300





● Molecule 1: Histone acetyltransferase p300



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	64.90Å 86.19Å 149.90Å 90.00° 96.98° 90.00°	Depositor
Resolution (Å)	32.30 – 2.33 32.30 – 2.33	Depositor EDS
% Data completeness (in resolution range)	91.8 (32.30-2.33) 99.1 (32.30-2.33)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.99 (at 2.34Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.229 , 0.275 0.227 , 0.222	Depositor DCC
$R_{free}$ test set	1982 reflections (2.85%)	DCC
Wilson B-factor (Å <sup>2</sup> )	35.5	Xtriage
Anisotropy	0.667	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 54.6	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	0.94	EDS
Total number of atoms	11399	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.96% 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: PEG, SO4, CL

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	0.40	0/959	0.57	1/1307 (0.1%)
1	B	0.34	0/940	0.50	0/1284
1	C	0.37	0/971	1.00	3/1322 (0.2%)
1	D	0.36	0/947	0.55	0/1291
1	E	0.38	0/975	0.55	1/1326 (0.1%)
1	F	0.36	0/917	0.51	1/1254 (0.1%)
1	G	0.39	0/945	0.54	0/1290
1	H	0.32	0/949	0.51	0/1296
1	I	0.41	0/971	1.05	3/1322 (0.2%)
1	J	0.34	0/952	0.52	0/1299
1	K	0.34	0/961	0.51	1/1310 (0.1%)
1	L	0.33	0/933	0.55	2/1274 (0.2%)
All	All	0.36	0/11420	0.64	12/15575 (0.1%)

There are no bond length outliers.

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	1137	ARG	NE-CZ-NH1	-21.85	109.37	120.30
1	C	1137	ARG	NE-CZ-NH1	-21.78	109.41	120.30
1	I	1137	ARG	NE-CZ-NH2	20.72	130.66	120.30
1	C	1137	ARG	NE-CZ-NH2	20.01	130.31	120.30
1	I	1137	ARG	CD-NE-CZ	10.35	138.09	123.60
1	C	1137	ARG	CD-NE-CZ	10.34	138.07	123.60
1	E	1137	ARG	NE-CZ-NH2	-6.51	117.05	120.30
1	L	1137	ARG	NE-CZ-NH2	-5.88	117.36	120.30
1	K	1137	ARG	NE-CZ-NH2	-5.50	117.55	120.30
1	A	1137	ARG	NE-CZ-NH2	-5.49	117.56	120.30
1	L	1137	ARG	NE-CZ-NH1	5.38	122.99	120.30
1	F	1137	ARG	NE-CZ-NH2	-5.07	117.77	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	932	0	889	28	0
1	B	913	0	856	16	0
1	C	944	0	910	16	0
1	D	920	0	874	23	0
1	E	948	0	914	23	0
1	F	890	0	828	16	0
1	G	918	0	871	11	0
1	H	922	0	867	19	0
1	I	944	0	907	19	0
1	J	925	0	881	17	0
1	K	934	0	892	28	0
1	L	906	0	849	5	0
2	A	1	0	0	1	0
2	E	1	0	0	1	0
3	A	7	0	10	1	0
3	E	7	0	10	0	0
4	E	5	0	0	0	0
4	L	5	0	0	1	0
5	A	39	0	0	1	0
5	B	14	0	0	0	0
5	C	31	0	0	1	0
5	D	20	0	0	0	0
5	E	32	0	0	0	0
5	F	1	0	0	0	0
5	G	38	0	0	0	0
5	H	15	0	0	0	0
5	I	24	0	0	0	0
5	J	23	0	0	1	0
5	K	28	0	0	0	0
5	L	12	0	0	0	0
All	All	11399	0	10558	166	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 8.

All (166) 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:1150:GLU:OE2	1:E:1140:LYS:HG2	1.17	1.28
1:A:1150:GLU:OE2	1:E:1140:LYS:CG	2.06	1.02
1:B:1115:TRP:HZ2	1:I:1135:THR:O	1.46	0.97
1:A:1150:GLU:OE1	1:E:1140:LYS:NZ	1.98	0.96
1:D:1134:LYS:HE2	1:K:1092:ILE:CD1	2.01	0.89
1:F:1134:LYS:HB3	1:F:1134:LYS:NZ	1.88	0.89
1:B:1137:ARG:HA	1:H:1115:TRP:CH2	2.10	0.86
1:D:1134:LYS:HE2	1:K:1092:ILE:HD13	1.56	0.85
1:B:1115:TRP:CZ2	1:I:1135:THR:O	2.29	0.85
1:C:1100:SER:O	1:C:1104:ARG:HG2	1.76	0.84
1:K:1134:LYS:NZ	1:K:1134:LYS:HB2	1.93	0.84
1:F:1090:PHE:HE2	1:J:1108:THR:O	1.62	0.83
1:F:1112:GLN:HE22	1:J:1090:PHE:HB3	1.46	0.78
1:F:1090:PHE:CE2	1:J:1108:THR:O	2.39	0.76
1:A:1150:GLU:OE2	1:E:1140:LYS:CE	2.37	0.71
1:A:1150:GLU:CD	1:E:1140:LYS:NZ	2.43	0.71
1:D:1084:LEU:HD22	5:J:384:HOH:O	1.89	0.70
1:A:1150:GLU:OE2	1:E:1140:LYS:HE3	1.96	0.66
1:I:1070:PRO:HG3	1:K:1137:ARG:CZ	2.27	0.65
1:A:1115:TRP:CZ2	1:G:1137:ARG:HA	2.31	0.65
1:F:1134:LYS:HZ1	1:F:1134:LYS:HB3	1.61	0.65
1:K:1134:LYS:HZ2	1:K:1134:LYS:HB2	1.61	0.65
1:D:1083:LEU:HD12	1:J:1063:GLU:HG2	1.80	0.64
1:D:1070:PRO:HG3	1:J:1137:ARG:NH2	2.14	0.63
1:I:1070:PRO:HG3	1:K:1137:ARG:NH2	2.15	0.62
1:D:1083:LEU:O	1:J:1067:ARG:NE	2.26	0.61
1:D:1134:LYS:HE2	1:K:1092:ILE:HD11	1.82	0.61
1:A:1143:SER:HB3	1:E:1135:THR:HG21	1.82	0.60
1:K:1134:LYS:HB2	1:K:1134:LYS:HZ3	1.64	0.60
1:A:1108:THR:HA	1:K:1082:GLN:HE22	1.68	0.59
2:E:2:CL:CL	1:G:1070:PRO:HB3	2.40	0.59
1:A:1077:GLN:NE2	5:A:290:HOH:O	2.25	0.59
1:H:1052:GLU:HA	1:H:1055[B]:ARG:NH1	2.19	0.58
1:B:1140:LYS:HE3	1:H:1116:GLN:NE2	2.17	0.58
1:A:1150:GLU:CD	1:E:1140:LYS:CE	2.74	0.57
1:E:1073:LEU:HB2	1:E:1074:PRO:HD3	1.87	0.56
1:D:1070:PRO:HG3	1:J:1137:ARG:CZ	2.36	0.56
1:I:1071(A):GLU:HG2	1:I:1144:LYS:HG2	1.88	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:1139:TYR:OH	1:K:1130:LEU:HD11	2.06	0.55
1:K:1134:LYS:NZ	1:K:1134:LYS:CB	2.65	0.53
1:A:1071(A):GLU:HG2	1:A:1144:LYS:HG2	1.91	0.53
1:C:1137:ARG:HD2	5:C:82:HOH:O	2.08	0.53
1:L:1071(A):GLU:HG2	1:L:1144:LYS:HG2	1.91	0.52
1:J:1071(A):GLU:HG2	1:J:1144:LYS:HG2	1.91	0.52
1:B:1140:LYS:HE3	1:H:1116:GLN:HE21	1.72	0.52
1:B:1071(A):GLU:HG2	1:B:1144:LYS:HG2	1.92	0.52
1:H:1071(A):GLU:HG2	1:H:1144:LYS:HG2	1.91	0.51
1:E:1136:SER:O	1:E:1140:LYS:HD3	2.10	0.51
1:A:1150:GLU:CD	1:E:1140:LYS:HE3	2.31	0.50
1:B:1150:GLU:OE2	1:K:1140:LYS:HE3	2.12	0.50
3:A:1:PEG:O2	1:E:1141:TYR:HE1	1.95	0.50
1:B:1137:ARG:N	1:H:1115:TRP:CZ2	2.80	0.50
1:G:1103:LYS:HE2	1:G:1107:ASP:OD2	2.11	0.50
1:I:1066:TYR:CD1	1:K:1083:LEU:HD13	2.47	0.50
1:F:1083:LEU:HD23	1:H:1063:GLU:HG2	1.93	0.50
1:A:1150:GLU:OE2	1:E:1140:LYS:CD	2.60	0.49
1:G:1071(A):GLU:HG2	1:G:1144:LYS:HG2	1.93	0.49
1:D:1103:LYS:HE2	1:D:1107:ASP:OD2	2.13	0.49
1:D:1137:ARG:CZ	1:J:1070:PRO:HG3	2.42	0.49
1:C:1073:LEU:HB2	1:C:1074:PRO:HD3	1.95	0.49
1:C:1136:SER:O	1:C:1140:LYS:HD3	2.12	0.49
1:J:1143:SER:O	1:J:1146:SER:HB3	2.13	0.49
1:I:1066:TYR:HD2	1:K:1084:LEU:HD23	1.79	0.48
1:H:1103:LYS:HE2	1:H:1107:ASP:OD2	2.13	0.48
1:A:1108:THR:HA	1:K:1082:GLN:NE2	2.28	0.48
1:F:1071(A):GLU:HG2	1:F:1144:LYS:HG2	1.95	0.48
1:F:1134:LYS:HB3	1:F:1134:LYS:HZ2	1.74	0.47
1:A:1106:LEU:HA	2:A:3:CL:CL	2.51	0.47
1:C:1082:GLN:HG2	1:C:1083:LEU:N	2.29	0.47
1:B:1112:GLN:HG3	1:I:1135:THR:HG21	1.96	0.47
1:K:1073:LEU:HB2	1:K:1074:PRO:HD3	1.96	0.47
1:L:1103:LYS:HE2	1:L:1107:ASP:OD2	2.15	0.47
1:E:1071(A):GLU:HG2	1:E:1144:LYS:HG2	1.97	0.47
1:J:1103:LYS:HE2	1:J:1107:ASP:OD2	2.13	0.47
1:C:1103:LYS:HE2	1:C:1107:ASP:OD2	2.15	0.47
1:H:1154:ASP:HB2	1:H:1155:PRO:HD3	1.96	0.47
1:I:1103:LYS:HZ1	1:K:1082:GLN:HE22	1.63	0.46
1:D:1134:LYS:CE	1:K:1092:ILE:HD13	2.39	0.46
1:B:1154:ASP:HB2	1:B:1155:PRO:HD3	1.96	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:1077:GLN:OE1	1:K:1077:GLN:OE1	2.34	0.46
1:G:1073:LEU:HB2	1:G:1074:PRO:HD3	1.98	0.46
1:B:1103:LYS:HE2	1:B:1107:ASP:OD2	2.16	0.46
1:C:1101:THR:HA	1:C:1104:ARG:HG3	1.97	0.46
1:K:1154:ASP:HB2	1:K:1155:PRO:HD3	1.98	0.46
1:K:1071(A):GLU:HG2	1:K:1144:LYS:HG2	1.98	0.46
1:D:1154:ASP:HB2	1:D:1155:PRO:HD3	1.97	0.45
1:J:1154:ASP:HB2	1:J:1155:PRO:HD3	1.97	0.45
1:F:1083:LEU:CD2	1:H:1063:GLU:HG2	2.47	0.45
1:F:1154:ASP:HB2	1:F:1155:PRO:HD3	1.98	0.45
1:K:1086:ILE:HG13	1:K:1086:ILE:O	2.16	0.45
1:L:1154:ASP:HB2	1:L:1155:PRO:HD3	1.98	0.45
1:C:1086:ILE:O	1:C:1086:ILE:HG13	2.16	0.45
1:B:1137:ARG:CA	1:H:1115:TRP:CH2	2.93	0.45
1:L:1133:ARG:HD3	4:L:2:SO4:O2	2.15	0.45
1:A:1086:ILE:O	1:A:1086:ILE:HG13	2.17	0.45
1:G:1154:ASP:HB2	1:G:1155:PRO:HD3	1.99	0.45
1:E:1103:LYS:HE2	1:E:1107:ASP:OD2	2.16	0.45
1:I:1137:ARG:CZ	1:K:1070:PRO:HG3	2.47	0.45
1:C:1071(A):GLU:HG2	1:C:1144:LYS:HG2	1.98	0.45
1:A:1134:LYS:HA	1:A:1139:TYR:CD1	2.51	0.45
1:E:1072:SER:HA	1:E:1145:LEU:HD11	1.99	0.45
1:F:1103:LYS:HE2	1:F:1107:ASP:OD2	2.16	0.45
1:C:1154:ASP:HB2	1:C:1155:PRO:HD3	1.99	0.44
1:I:1066:TYR:CE1	1:K:1083:LEU:HD13	2.52	0.44
1:B:1137:ARG:CA	1:H:1115:TRP:CZ2	3.01	0.44
1:D:1086:ILE:O	1:D:1086:ILE:HG13	2.18	0.44
1:K:1103:LYS:HE2	1:K:1107:ASP:OD2	2.17	0.44
1:F:1086:ILE:O	1:F:1086:ILE:HG13	2.18	0.44
1:F:1072:SER:HA	1:F:1145:LEU:HD11	2.00	0.44
1:A:1129:TRP:CH2	1:E:1135:THR:O	2.71	0.44
1:G:1094:LYS:HA	1:G:1094:LYS:HD3	1.79	0.44
1:A:1154:ASP:HB2	1:A:1155:PRO:HD3	1.99	0.43
1:J:1072:SER:HA	1:J:1145:LEU:HD11	2.00	0.43
1:A:1153:ILE:HG13	1:A:1157:MET:HG2	1.99	0.43
1:I:1153:ILE:HG13	1:I:1157:MET:HG2	2.01	0.43
1:E:1154:ASP:HB2	1:E:1155:PRO:HD3	1.99	0.43
1:D:1071(A):GLU:HG2	1:D:1144:LYS:HG2	2.00	0.43
1:B:1137:ARG:HA	1:H:1115:TRP:CZ2	2.50	0.43
1:H:1072:SER:HA	1:H:1145:LEU:HD11	2.01	0.43
1:A:1073:LEU:HB2	1:A:1074:PRO:HD3	2.01	0.43

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:1072:SER:HA	1:C:1145:LEU:HD11	2.01	0.42
1:I:1154:ASP:HB2	1:I:1155:PRO:HD3	2.01	0.42
1:D:1153:ILE:HG13	1:D:1157:MET:HG2	2.01	0.42
1:D:1123:LEU:HA	1:K:1134:LYS:HE2	2.01	0.42
1:L:1086:ILE:O	1:L:1086:ILE:HG13	2.19	0.42
1:C:1101:THR:HA	1:C:1104:ARG:CG	2.50	0.42
1:D:1073:LEU:HB2	1:D:1074:PRO:HD3	2.01	0.42
1:C:1050:LYS:HA	1:C:1051:PRO:HD3	1.94	0.42
1:I:1134:LYS:HA	1:I:1139:TYR:CD1	2.54	0.42
1:C:1153:ILE:HG13	1:C:1157:MET:HG2	2.02	0.42
1:A:1134:LYS:HA	1:A:1139:TYR:CG	2.55	0.42
1:F:1073:LEU:HB2	1:F:1074:PRO:HD3	2.01	0.42
1:F:1073:LEU:HB3	1:H:1073:LEU:HB3	2.02	0.42
1:A:1103:LYS:HE2	1:A:1107:ASP:OD2	2.19	0.42
1:F:1153:ILE:HG13	1:F:1157:MET:HG2	2.02	0.42
1:I:1086:ILE:O	1:I:1086:ILE:HG13	2.19	0.42
1:E:1094:LYS:HD3	1:E:1094:LYS:HA	1.79	0.42
1:A:1136:SER:O	1:A:1140:LYS:HD3	2.20	0.41
1:A:1072:SER:HA	1:A:1145:LEU:HD11	2.00	0.41
1:C:1094:LYS:HA	1:C:1094:LYS:HD3	1.78	0.41
1:D:1137:ARG:NH2	1:J:1070:PRO:HG3	2.35	0.41
1:G:1153:ILE:HG13	1:G:1157:MET:HG2	2.02	0.41
1:B:1153:ILE:HG13	1:B:1157:MET:HG2	2.02	0.41
1:I:1072:SER:HA	1:I:1145:LEU:HD11	2.01	0.41
1:A:1073:LEU:HB3	1:C:1073:LEU:HB3	2.02	0.41
1:G:1136:SER:O	1:G:1140:LYS:HD3	2.21	0.41
1:H:1153:ILE:HG13	1:H:1157:MET:HG2	2.01	0.41
1:I:1094:LYS:HD3	1:I:1094:LYS:HA	1.86	0.41
1:B:1073:LEU:HB2	1:B:1074:PRO:HD3	2.03	0.41
1:D:1136:SER:O	1:D:1140:LYS:HD3	2.19	0.41
1:E:1153:ILE:HG13	1:E:1157:MET:HG2	2.02	0.41
1:J:1153:ILE:HG13	1:J:1157:MET:HG2	2.02	0.41
1:E:1086:ILE:O	1:E:1086:ILE:HG13	2.21	0.41
1:H:1086:ILE:O	1:H:1086:ILE:HG13	2.20	0.41
1:H:1134:LYS:HA	1:H:1139:TYR:CD1	2.56	0.41
1:G:1134:LYS:HA	1:G:1139:TYR:CD1	2.56	0.41
1:K:1153:ILE:HG13	1:K:1157:MET:HG2	2.02	0.41
1:H:1073:LEU:HB2	1:H:1074:PRO:HD3	2.03	0.40
1:D:1134:LYS:HA	1:D:1139:TYR:CD1	2.56	0.40
1:E:1083:LEU:HD13	1:G:1066:TYR:CD1	2.57	0.40
1:A:1059:MET:N	1:A:1060:PRO:CD	2.85	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:1134:LYS:HA	1:E:1139:TYR:CD1	2.56	0.40
1:I:1103:LYS:NZ	1:K:1082:GLN:NE2	2.70	0.40
1:D:1083:LEU:O	1:J:1067:ARG:CG	2.69	0.40
1:J:1094:LYS:HD3	1:J:1094:LYS:HA	1.84	0.40
1:D:1050:LYS:HA	1:D:1051:PRO:HD3	1.95	0.40

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	112/124 (90%)	112 (100%)	0	0	100	100
1	B	112/124 (90%)	112 (100%)	0	0	100	100
1	C	112/124 (90%)	112 (100%)	0	0	100	100
1	D	111/124 (90%)	110 (99%)	1 (1%)	0	100	100
1	E	113/124 (91%)	113 (100%)	0	0	100	100
1	F	109/124 (88%)	109 (100%)	0	0	100	100
1	G	111/124 (90%)	111 (100%)	0	0	100	100
1	H	112/124 (90%)	112 (100%)	0	0	100	100
1	I	113/124 (91%)	113 (100%)	0	0	100	100
1	J	111/124 (90%)	111 (100%)	0	0	100	100
1	K	113/124 (91%)	113 (100%)	0	0	100	100
1	L	111/124 (90%)	111 (100%)	0	0	100	100
All	All	1340/1488 (90%)	1339 (100%)	1 (0%)	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	100/116 (86%)	100 (100%)	0	100	100
1	B	96/116 (83%)	95 (99%)	1 (1%)	80	88
1	C	104/116 (90%)	102 (98%)	2 (2%)	62	74
1	D	98/116 (84%)	96 (98%)	2 (2%)	60	73
1	E	104/116 (90%)	103 (99%)	1 (1%)	80	88
1	F	94/116 (81%)	93 (99%)	1 (1%)	78	87
1	G	99/116 (85%)	99 (100%)	0	100	100
1	H	97/116 (84%)	96 (99%)	1 (1%)	80	88
1	I	102/116 (88%)	99 (97%)	3 (3%)	48	58
1	J	100/116 (86%)	100 (100%)	0	100	100
1	K	99/116 (85%)	97 (98%)	2 (2%)	60	73
1	L	94/116 (81%)	93 (99%)	1 (1%)	78	87
All	All	1187/1392 (85%)	1173 (99%)	14 (1%)	75	86

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

Mol	Chain	Res	Type
1	B	1140	LYS
1	C	1104	ARG
1	C	1140	LYS
1	D	1083	LEU
1	D	1134	LYS
1	E	1053	GLU
1	F	1134	LYS
1	H	1083	LEU
1	I	1053	GLU
1	I	1083	LEU
1	I	1106	LEU
1	K	1134	LYS
1	K	1140	LYS
1	L	1083	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1082	GLN
1	C	1082	GLN
1	D	1126	ASN
1	F	1112	GLN
1	G	1082	GLN
1	K	1082	GLN
1	K	1126	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 2 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	$\# Z  > 2$
3	PEG	A	1	-	6,6,6	0.61	0	5,5,5	2.03	3 (60%)
4	SO4	E	1	-	4,4,4	0.25	0	6,6,6	0.19	0
3	PEG	E	1162	-	6,6,6	0.50	0	5,5,5	1.46	0
4	SO4	L	2	-	4,4,4	0.19	0	6,6,6	0.13	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
3	PEG	A	1	-	-	0/4/4/4	0/0/0/0
4	SO4	E	1	-	-	0/0/0/0	0/0/0/0
3	PEG	E	1162	-	-	0/4/4/4	0/0/0/0
4	SO4	L	2	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1	PEG	O2-C2-C1	2.23	120.43	110.15
3	A	1	PEG	C3-O2-C2	2.39	123.64	113.30
3	A	1	PEG	O2-C3-C4	2.61	122.21	110.15

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	A	1	PEG	1	0
4	L	2	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	114/124 (91%)	0.56	8 (7%) 17 25	23, 41, 90, 121	0
1	B	114/124 (91%)	0.40	8 (7%) 17 25	40, 61, 101, 120	0
1	C	114/124 (91%)	0.33	7 (6%) 22 31	25, 45, 88, 108	0
1	D	113/124 (91%)	0.37	7 (6%) 21 30	30, 51, 95, 114	0
1	E	115/124 (92%)	0.22	4 (3%) 44 56	23, 44, 86, 102	0
1	F	111/124 (89%)	1.39	30 (27%) 1 1	49, 88, 135, 178	0
1	G	113/124 (91%)	0.47	9 (7%) 13 19	24, 45, 92, 102	0
1	H	113/124 (91%)	0.59	8 (7%) 17 24	36, 64, 96, 111	0
1	I	115/124 (92%)	0.34	9 (7%) 14 20	22, 44, 81, 112	0
1	J	113/124 (91%)	0.45	11 (9%) 8 13	36, 52, 95, 100	0
1	K	114/124 (91%)	0.22	6 (5%) 27 38	29, 55, 101, 144	0
1	L	113/124 (91%)	0.78	12 (10%) 7 11	39, 70, 116, 145	0
All	All	1362/1488 (91%)	0.51	119 (8%) 11 16	22, 55, 102, 178	0

All (119) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	1160	LEU	11.9
1	J	1048	ILE	6.9
1	B	1161	GLY	6.5
1	C	1048	ILE	6.1
1	D	1160	LEU	6.1
1	F	1066	TYR	6.0
1	F	1159	SER	5.9
1	K	1050	LYS	5.7
1	I	1161	GLY	5.7
1	L	1049	PHE	5.6
1	A	1048	ILE	5.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	1057	ALA	5.4
1	F	1156	VAL	5.3
1	F	1118	VAL	5.1
1	B	1049	PHE	5.0
1	J	1160	LEU	4.6
1	F	1153	ILE	4.6
1	F	1049	PHE	4.4
1	L	1159	SER	4.4
1	L	1155	PRO	4.3
1	L	1048	ILE	4.2
1	L	1156	VAL	4.2
1	J	1049	PHE	4.2
1	G	1160	LEU	4.2
1	J	1156	VAL	4.2
1	H	1158	GLN	4.1
1	A	1092	ILE	3.9
1	D	1049	PHE	3.9
1	F	1149	PHE	3.8
1	L	1157	MET	3.8
1	A	1049	PHE	3.8
1	E	1048	ILE	3.8
1	A	1085	GLY	3.7
1	F	1157	MET	3.7
1	F	1115	TRP	3.7
1	H	1049	PHE	3.6
1	C	1160	LEU	3.6
1	F	1138	VAL	3.5
1	F	1053	GLU	3.5
1	H	1066	TYR	3.5
1	E	1049	PHE	3.4
1	J	1066	TYR	3.4
1	F	1054	LEU	3.4
1	F	1061	THR	3.4
1	I	1155	PRO	3.4
1	A	1160	LEU	3.3
1	F	1158	GLN	3.2
1	L	1051	PRO	3.1
1	C	1115	TRP	3.1
1	F	1112	GLN	3.1
1	F	1113	GLU	3.0
1	K	1114	PRO	3.0
1	D	1048	ILE	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	1049	PHE	2.9
1	G	1087	PRO	2.8
1	F	1106	LEU	2.7
1	H	1159	SER	2.7
1	B	1160	LEU	2.7
1	F	1062	LEU	2.7
1	F	1055	ARG	2.7
1	L	1113	GLU	2.7
1	G	1055	ARG	2.7
1	C	1047	LYS	2.7
1	J	1051	PRO	2.7
1	F	1051	PRO	2.6
1	H	1156	VAL	2.6
1	F	1114	PRO	2.6
1	D	1158	GLN	2.6
1	I	1049	PHE	2.6
1	B	1087	PRO	2.5
1	H	1152	GLU	2.5
1	J	1158	GLN	2.5
1	F	1109	GLY	2.5
1	K	1049	PHE	2.5
1	F	1111	TYR	2.5
1	L	1050	LYS	2.5
1	F	1154	ASP	2.5
1	E	1051	PRO	2.5
1	C	1094	LYS	2.5
1	E	1134	LYS	2.5
1	F	1155	PRO	2.4
1	I	1048	ILE	2.4
1	H	1070	PRO	2.4
1	B	1048	ILE	2.4
1	G	1049	PHE	2.4
1	J	1159	SER	2.4
1	F	1099	LEU	2.3
1	K	1115	TRP	2.3
1	K	1054	LEU	2.3
1	L	1137	ARG	2.3
1	J	1151	GLN	2.3
1	D	1054	LEU	2.2
1	I	1050	LYS	2.2
1	G	1048	ILE	2.2
1	K	1048	ILE	2.2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	1154	ASP	2.2
1	B	1051	PRO	2.2
1	J	1054	LEU	2.2
1	C	1137	ARG	2.2
1	H	1048	ILE	2.2
1	F	1130	LEU	2.2
1	D	1050	LYS	2.1
1	F	1050	LYS	2.1
1	F	1141	TYR	2.1
1	A	1147	GLU	2.1
1	D	1083	LEU	2.1
1	G	1158	GLN	2.1
1	I	1085	GLY	2.1
1	A	1091	ASP	2.1
1	B	1054	LEU	2.1
1	I	1159	SER	2.1
1	L	1115	TRP	2.1
1	G	1156	VAL	2.1
1	G	1152	GLU	2.1
1	I	1158	GLN	2.1
1	G	1085	GLY	2.1
1	I	1092	ILE	2.0
1	J	1152	GLU	2.0
1	A	1108	THR	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
3	PEG	A	1	7/7	0.72	0.40	4.46	48,48,48,48	0
2	CL	A	3	1/1	0.97	0.26	2.66	61,61,61,61	0
2	CL	E	2	1/1	0.92	0.13	-1.14	67,67,67,67	0
3	PEG	E	1162	7/7	0.86	0.24	-	66,66,66,66	0
4	SO4	L	2	5/5	0.95	0.18	-	52,57,109,127	0
4	SO4	E	1	5/5	0.91	0.12	-	42,49,65,96	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.