



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 12:35 am BST

PDB ID : 5M3H  
Title : Bat influenza A/H17N10 polymerase bound to four heptad repeats of serine 5 phosphorylated Pol II CTD  
Authors : Lukarska, M.; Pflug, A.; Cusack, S.  
Deposited on : 2016-10-14  
Resolution : 2.50 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

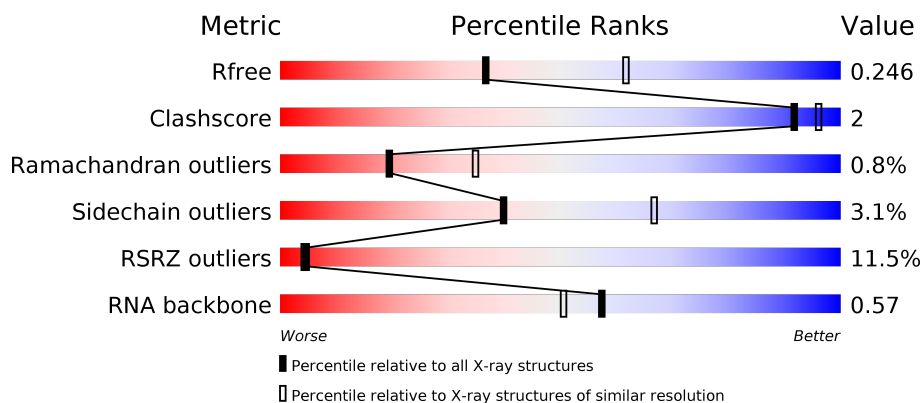
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)
RNA backbone	3102	1008 (2.84-2.16)

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	728	<div> <div>21%</div> <div> <div></div> <div>88%</div> <div>7%</div> <div>••</div> </div> </div>
2	B	776	<div> <div>4%</div> <div> <div></div> <div>91%</div> <div>5%</div> <div>•</div> </div> </div>
3	C	797	<div> <div>9%</div> <div> <div></div> <div>86%</div> <div>5%</div> <div>8%</div> </div> </div>
4	R	18	<div> <div></div> <div> <div></div> <div>61%</div> <div>11%</div> <div>28%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	V	16	<div><div></div><div>6%</div><div>75%</div><div>19%</div><div>6%</div></div>
6	X	28	<div><div></div><div>7%</div><div>18%</div><div>79%</div></div>
6	Y	28	<div><div></div><div>36%</div><div>64%</div></div>

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 18529 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 Polymerase acidic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	699	Total	C	N	O	S	0	0	0
			5692	3617	959	1079	37			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP H6QM92
A	-2	SER	-	expression tag	UNP H6QM92
A	-1	GLY	-	expression tag	UNP H6QM92
A	0	SER	-	expression tag	UNP H6QM92
A	714	GLY	-	expression tag	UNP H6QM92
A	715	SER	-	expression tag	UNP H6QM92
A	716	GLY	-	expression tag	UNP H6QM92
A	717	SER	-	expression tag	UNP H6QM92
A	718	GLY	-	expression tag	UNP H6QM92
A	719	GLU	-	expression tag	UNP H6QM92
A	720	ASN	-	expression tag	UNP H6QM92
A	721	LEU	-	expression tag	UNP H6QM92
A	722	TYR	-	expression tag	UNP H6QM92
A	723	PHE	-	expression tag	UNP H6QM92
A	724	GLN	-	expression tag	UNP H6QM92

- Molecule 2 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	750	Total	C	N	O	S	0	0	0
			5986	3766	1062	1118	40			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	GLY	-	expression tag	UNP H6QM91

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-7	SER	-	expression tag	UNP H6QM91
B	-6	GLY	-	expression tag	UNP H6QM91
B	-5	SER	-	expression tag	UNP H6QM91
B	-4	GLY	-	expression tag	UNP H6QM91
B	-3	SER	-	expression tag	UNP H6QM91
B	-2	GLY	-	expression tag	UNP H6QM91
B	-1	SER	-	expression tag	UNP H6QM91
B	0	GLY	-	expression tag	UNP H6QM91
B	757	GLY	-	expression tag	UNP H6QM91
B	758	SER	-	expression tag	UNP H6QM91
B	759	GLY	-	expression tag	UNP H6QM91
B	760	SER	-	expression tag	UNP H6QM91
B	761	GLY	-	expression tag	UNP H6QM91
B	762	GLU	-	expression tag	UNP H6QM91
B	763	ASN	-	expression tag	UNP H6QM91
B	764	LEU	-	expression tag	UNP H6QM91
B	765	TYR	-	expression tag	UNP H6QM91
B	766	PHE	-	expression tag	UNP H6QM91
B	767	GLN	-	expression tag	UNP H6QM91

- Molecule 3 is a protein called Polymerase basic protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	730	Total	C	N	O	S	0	0	0
			5796	3657	1025	1082	32			

There are 37 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	GLY	-	expression tag	UNP H6QM90
C	-7	SER	-	expression tag	UNP H6QM90
C	-6	GLY	-	expression tag	UNP H6QM90
C	-5	SER	-	expression tag	UNP H6QM90
C	-4	GLY	-	expression tag	UNP H6QM90
C	-3	SER	-	expression tag	UNP H6QM90
C	-2	GLY	-	expression tag	UNP H6QM90
C	-1	SER	-	expression tag	UNP H6QM90
C	0	GLY	-	expression tag	UNP H6QM90
C	761	GLY	-	expression tag	UNP H6QM90
C	762	TRP	-	expression tag	UNP H6QM90
C	763	SER	-	expression tag	UNP H6QM90
C	764	HIS	-	expression tag	UNP H6QM90

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	765	PRO	-	expression tag	UNP H6QM90
C	766	GLN	-	expression tag	UNP H6QM90
C	767	PHE	-	expression tag	UNP H6QM90
C	768	GLU	-	expression tag	UNP H6QM90
C	769	LYS	-	expression tag	UNP H6QM90
C	770	GLY	-	expression tag	UNP H6QM90
C	771	TRP	-	expression tag	UNP H6QM90
C	772	SER	-	expression tag	UNP H6QM90
C	773	HIS	-	expression tag	UNP H6QM90
C	774	PRO	-	expression tag	UNP H6QM90
C	775	GLN	-	expression tag	UNP H6QM90
C	776	PHE	-	expression tag	UNP H6QM90
C	777	GLU	-	expression tag	UNP H6QM90
C	778	LYS	-	expression tag	UNP H6QM90
C	779	GLY	-	expression tag	UNP H6QM90
C	780	SER	-	expression tag	UNP H6QM90
C	781	GLY	-	expression tag	UNP H6QM90
C	782	SER	-	expression tag	UNP H6QM90
C	783	GLU	-	expression tag	UNP H6QM90
C	784	ASN	-	expression tag	UNP H6QM90
C	785	LEU	-	expression tag	UNP H6QM90
C	786	TYR	-	expression tag	UNP H6QM90
C	787	PHE	-	expression tag	UNP H6QM90
C	788	GLN	-	expression tag	UNP H6QM90

- Molecule 4 is a RNA chain called RNA (5'-R(\*UP\*AP\*UP\*AP\*CP\*CP\*UP\*CP\*UP\*GP\*CP\*UP\*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	R	13	Total	C	N	O	P	0	0	0
			247	111	37	87	12			

- Molecule 5 is a RNA chain called RNA 5'-pAGUAGUAACAAGAGGG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	V	16	Total	C	N	O	P	0	0	0
			353	157	72	108	16			

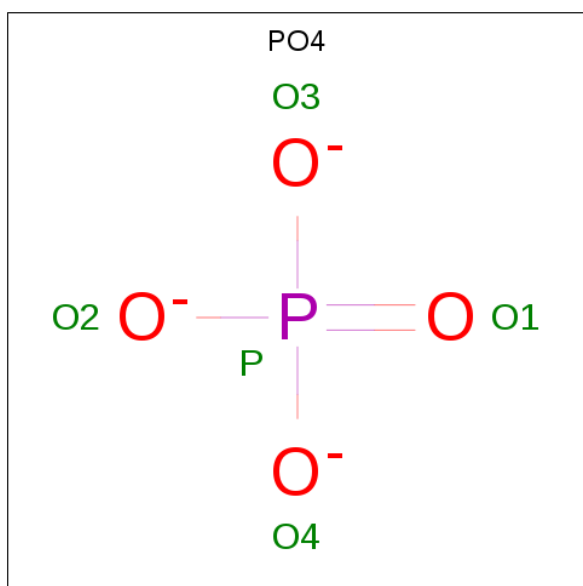
- Molecule 6 is a protein called TYR-SER-PRO-THR-SEP-PRO.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	X	6	Total	C	N	O	P	0	0	0
			49	29	6	13	1			
6	Y	10	Total	C	N	O	P	0	0	0
			80	49	10	20	1			

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Zn	0	0
			1	1		

- Molecule 8 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total	O	P	0	0
			5	4	1		
8	B	1	Total	O	P	0	0
			5	4	1		
8	B	1	Total	O	P	0	0
			5	4	1		
8	C	1	Total	O	P	0	0
			5	4	1		
8	C	1	Total	O	P	0	0
			5	4	1		
8	C	1	Total	O	P	0	0
			5	4	1		
8	C	1	Total	O	P	0	0
			5	4	1		

- Molecule 9 is water.

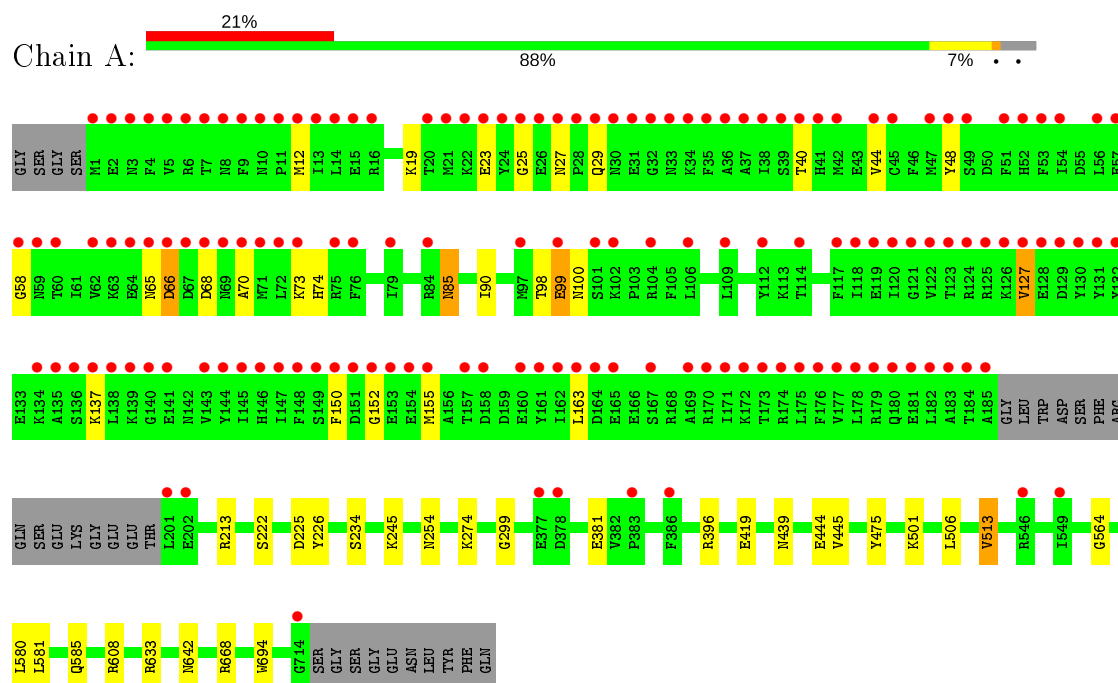
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	78	Total	O	0	0
			78	78		
9	B	105	Total	O	0	0
			105	105		
9	C	56	Total	O	0	0
			56	56		
9	R	3	Total	O	0	0
			3	3		
9	V	7	Total	O	0	0
			7	7		
9	X	1	Total	O	0	0
			1	1		



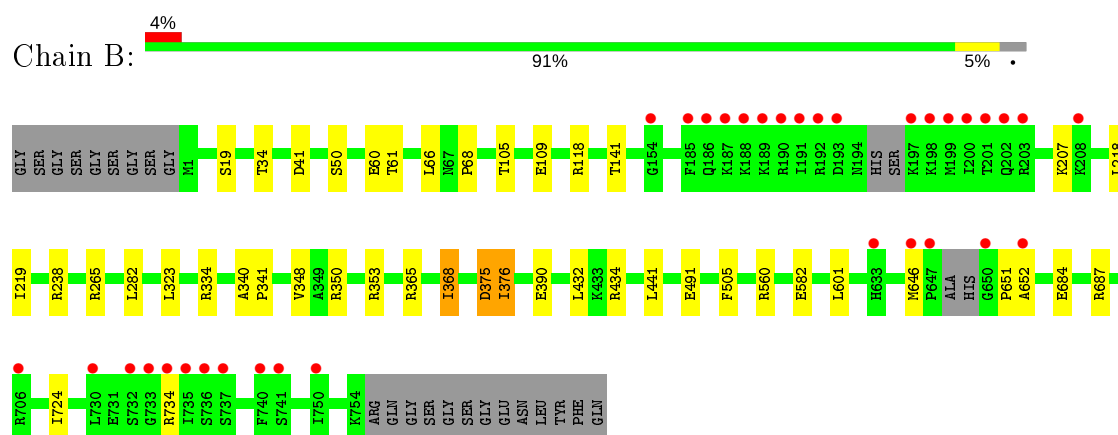
### 3 Residue-property plots [i](#)

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: Polymerase acidic protein



#### • Molecule 2: RNA-directed RNA polymerase catalytic subunit



#### • Molecule 3: Polymerase basic protein 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	269.49 Å   147.52 Å   88.48 Å 90.00°   97.22°   90.00°	Depositor
Resolution (Å)	49.53 – 2.50 49.53 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.53-2.50) 99.6 (49.53-2.50)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.58 (at 2.51 Å)	Xtriage
Refinement program	PHENIX (1.10 _2155: ???)	Depositor
R, $R_{free}$	0.210   ,   0.245 0.213   ,   0.246	Depositor DCC
$R_{free}$ test set	5757 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	55.2	Xtriage
Anisotropy	0.219	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 54.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	18529	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.91% 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: PO4, ZN, SEP

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	0.25	0/5812	0.42	0/7828
2	B	0.24	0/6100	0.48	3/8229 (0.0%)
3	C	0.24	0/5894	0.43	0/7955
4	R	0.23	0/273	0.74	0/421
5	V	0.56	1/397 (0.3%)	0.73	0/617
6	X	0.33	0/40	0.38	0/53
6	Y	0.22	0/73	0.37	0/98
All	All	0.25	1/18589 (0.0%)	0.46	3/25201 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	V	1	A	OP3-P	-10.53	1.48	1.61

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	376	ILE	N-CA-C	11.62	142.38	111.00
2	B	375	ASP	CB-CA-C	9.40	129.20	110.40
2	B	376	ILE	N-CA-CB	-6.42	96.04	110.80

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	5692	0	5613	17	0
2	B	5986	0	6006	22	0
3	C	5796	0	5923	21	0
4	R	247	0	128	4	0
5	V	353	0	175	1	0
6	X	49	0	40	1	0
6	Y	80	0	64	0	0
7	A	1	0	0	0	0
8	A	15	0	0	0	0
8	B	35	0	0	0	0
8	C	25	0	0	1	0
9	A	78	0	0	0	0
9	B	105	0	0	0	0
9	C	56	0	0	1	0
9	R	3	0	0	0	0
9	V	7	0	0	0	0
9	X	1	0	0	0	0
All	All	18529	0	17949	59	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 (59) 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)
2:B:61:THR:HG22	2:B:218:LEU:HD23	1.36	1.05
4:R:4:A:H5''	4:R:4:A:H8	1.36	0.88
2:B:61:THR:CG2	2:B:218:LEU:HD23	2.06	0.84
4:R:4:A:H5''	4:R:4:A:C8	2.19	0.77
3:C:286:THR:HB	3:C:532:ASN:O	1.86	0.74
3:C:286:THR:O	3:C:532:ASN:ND2	2.19	0.73
2:B:61:THR:CG2	2:B:218:LEU:CD2	2.70	0.69
3:C:144:ARG:NH2	8:C:801:PO4:O1	2.28	0.67
2:B:61:THR:HG22	2:B:218:LEU:CD2	2.18	0.65
4:R:4:A:H2'	4:R:5:C:O4'	2.02	0.59
3:C:530:ASN:ND2	3:C:532:ASN:OD1	2.40	0.55
2:B:368:ILE:O	2:B:368:ILE:HG13	2.06	0.55
1:A:213:ARG:NH2	2:B:60:GLU:OE2	2.40	0.53
1:A:444:GLU:OE2	1:A:633:ARG:NH1	2.41	0.52
1:A:585:GLN:HG2	2:B:505:PHE:CD2	2.44	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:34:THR:HG23	5:V:8:A:OP1	2.11	0.51
1:A:506:LEU:HD11	1:A:513:VAL:HG22	1.93	0.50
1:A:85:ASN:N	1:A:85:ASN:OD1	2.44	0.50
2:B:34:THR:HG22	2:B:353:ARG:HB2	1.94	0.49
3:C:290:SER:O	3:C:291:SER:OG	2.24	0.49
2:B:109:GLU:OE2	2:B:265:ARG:NH2	2.47	0.48
2:B:684:GLU:OE1	2:B:687:ARG:NH1	2.47	0.48
3:C:395:LEU:HD11	3:C:463:ILE:CD1	2.44	0.47
3:C:530:ASN:CG	3:C:532:ASN:OD1	2.53	0.47
3:C:57:ILE:O	3:C:91:VAL:N	2.48	0.46
1:A:98:THR:O	1:A:100:ASN:N	2.49	0.45
2:B:61:THR:HG23	2:B:218:LEU:CD2	2.44	0.45
3:C:144:ARG:NH1	9:C:903:HOH:O	2.48	0.45
2:B:141:THR:HG21	3:C:35:THR:O	2.17	0.45
4:R:4:A:C5'	4:R:4:A:C8	2.96	0.45
2:B:323:LEU:O	2:B:334:ARG:NH1	2.51	0.44
1:A:633:ARG:NH2	6:X:5:SEP:O2P	2.49	0.44
3:C:429:ASN:HB2	3:C:430:PRO:HD2	2.00	0.44
3:C:145:VAL:HG21	3:C:220:ILE:HD11	1.99	0.43
1:A:225:ASP:OD1	1:A:226:TYR:N	2.52	0.43
1:A:513:VAL:O	1:A:564:GLY:HA2	2.19	0.43
3:C:151:HIS:ND1	3:C:151:HIS:O	2.52	0.43
1:A:642:ASN:OD1	1:A:694:TRP:NE1	2.46	0.42
3:C:530:ASN:HD22	3:C:530:ASN:C	2.20	0.42
1:A:445:VAL:CG1	1:A:580:LEU:HD22	2.49	0.42
1:A:475:TYR:HA	1:A:501:LYS:O	2.19	0.42
3:C:399:ILE:HD13	3:C:402:MET:CE	2.50	0.42
3:C:462:GLY:HA3	3:C:474:THR:HB	2.02	0.42
3:C:686:ILE:HA	3:C:696:ILE:O	2.20	0.42
2:B:282:LEU:HG	2:B:441:LEU:HD22	2.01	0.42
1:A:65:ASN:O	1:A:66:ASP:CB	2.67	0.41
1:A:90:ILE:HA	2:B:724:ILE:HD13	2.01	0.41
2:B:34:THR:HG22	2:B:353:ARG:CB	2.51	0.41
3:C:429:ASN:HB2	3:C:430:PRO:CD	2.50	0.41
2:B:50:SER:HB3	2:B:68:PRO:HB3	2.03	0.41
1:A:274:LYS:O	1:A:396:ARG:HD2	2.20	0.41
1:A:234:SER:OG	2:B:334:ARG:NH2	2.54	0.41
2:B:61:THR:HB	2:B:219:ILE:HD11	2.03	0.41
3:C:344:VAL:HG12	3:C:345:LEU:N	2.36	0.41
1:A:40:THR:HG22	1:A:44:VAL:HG23	2.03	0.41
3:C:460:ILE:HD13	3:C:481:SER:HB2	2.03	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:286:THR:CB	3:C:532:ASN:O	2.64	0.40
2:B:340:ALA:HB3	2:B:341:PRO:HD3	2.03	0.40
2:B:432:LEU:HB3	2:B:434:ARG:HG2	2.03	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	695/728 (96%)	638 (92%)	45 (6%)	12 (2%)	9	16
2	B	744/776 (96%)	712 (96%)	28 (4%)	4 (0%)	29	48
3	C	724/797 (91%)	702 (97%)	21 (3%)	1 (0%)	51	73
6	X	3/28 (11%)	3 (100%)	0	0	100	100
6	Y	7/28 (25%)	7 (100%)	0	0	100	100
All	All	2173/2357 (92%)	2062 (95%)	94 (4%)	17 (1%)	19	35

All (17) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	376	ILE
1	A	66	ASP
1	A	68	ASP
1	A	99	GLU
1	A	19	LYS
1	A	25	GLY
1	A	137	LYS
1	A	152	GLY
2	B	375	ASP
2	B	652	ALA
1	A	23	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	70	ALA
2	B	651	PRO
1	A	127	VAL
3	C	152	LYS
1	A	58	GLY
1	A	299	GLY

### 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	625/648 (96%)	603 (96%)	22 (4%)	36	62
2	B	659/676 (98%)	641 (97%)	18 (3%)	44	71
3	C	647/703 (92%)	627 (97%)	20 (3%)	40	67
6	X	5/24 (21%)	5 (100%)	0	100	100
6	Y	9/24 (38%)	9 (100%)	0	100	100
All	All	1945/2075 (94%)	1885 (97%)	60 (3%)	40	67

All (60) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	MET
1	A	27	ASN
1	A	29	GLN
1	A	48	TYR
1	A	73	LYS
1	A	74	HIS
1	A	85	ASN
1	A	99	GLU
1	A	127	VAL
1	A	150	PHE
1	A	155	MET
1	A	163	LEU
1	A	222	SER

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	245	LYS
1	A	254	ASN
1	A	381	GLU
1	A	419	GLU
1	A	439	ASN
1	A	513	VAL
1	A	581	LEU
1	A	608	ARG
1	A	668	ARG
2	B	19	SER
2	B	41	ASP
2	B	66	LEU
2	B	105	THR
2	B	118	ARG
2	B	207	LYS
2	B	238	ARG
2	B	348	VAL
2	B	350	ARG
2	B	365	ARG
2	B	368	ILE
2	B	390	GLU
2	B	491	GLU
2	B	560	ARG
2	B	582	GLU
2	B	601	LEU
2	B	646	MET
2	B	734	ARG
3	C	15	ARG
3	C	65	GLU
3	C	87	ASN
3	C	164	MET
3	C	167	VAL
3	C	179	SER
3	C	258	THR
3	C	317	ILE
3	C	334	LYS
3	C	354	MET
3	C	358	GLU
3	C	392	GLU
3	C	410	MET
3	C	451	THR
3	C	530	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	532	ASN
3	C	536	MET
3	C	539	VAL
3	C	612	SER
3	C	678	SER

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

Mol	Chain	Res	Type
3	C	530	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	R	11/18 (61%)	2 (18%)	1 (9%)
5	V	15/16 (93%)	3 (20%)	0
All	All	26/34 (76%)	5 (19%)	1 (3%)

All (5) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	R	4	A
4	R	5	C
5	V	7	A
5	V	8	A
5	V	11	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	R	4	A

## 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	$\# Z  > 2$
6	SEP	X	5	6	8,9,10	1.55	1 (12%)	8,12,14	1.40	2 (25%)
6	SEP	Y	19	6	8,9,10	1.57	1 (12%)	8,12,14	1.39	2 (25%)

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	SEP	X	5	6	-	2/5/8/10	-
6	SEP	Y	19	6	-	3/5/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	Y	19	SEP	P-O1P	3.44	1.61	1.50
6	X	5	SEP	P-O1P	3.39	1.61	1.50

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	Y	19	SEP	OG-CB-CA	2.77	110.84	108.14
6	X	5	SEP	P-OG-CB	-2.57	111.22	118.30
6	X	5	SEP	OG-CB-CA	2.49	110.57	108.14
6	Y	19	SEP	O2P-P-OG	2.11	112.34	106.73

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	Y	19	SEP	CB-OG-P-O1P
6	Y	19	SEP	CB-OG-P-O2P
6	Y	19	SEP	CB-OG-P-O3P
6	X	5	SEP	N-CA-CB-OG
6	X	5	SEP	CA-CB-OG-P

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	X	5	SEP	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, 1 is monoatomic - leaving 15 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	$\# Z  > 2$
8	PO4	A	803	-	4,4,4	0.90	0	6,6,6	0.44	0
8	PO4	B	805	-	4,4,4	0.90	0	6,6,6	0.46	0
8	PO4	C	803	-	4,4,4	0.90	0	6,6,6	0.45	0
8	PO4	A	804	-	4,4,4	0.91	0	6,6,6	0.43	0
8	PO4	B	804	-	4,4,4	0.93	0	6,6,6	0.41	0
8	PO4	C	804	-	4,4,4	0.91	0	6,6,6	0.43	0
8	PO4	B	807	-	4,4,4	0.92	0	6,6,6	0.43	0
8	PO4	C	805	-	4,4,4	0.93	0	6,6,6	0.43	0
8	PO4	A	802	-	4,4,4	0.90	0	6,6,6	0.44	0
8	PO4	B	803	-	4,4,4	0.90	0	6,6,6	0.44	0
8	PO4	C	801	-	4,4,4	0.90	0	6,6,6	0.46	0
8	PO4	B	801	-	4,4,4	0.90	0	6,6,6	0.44	0
8	PO4	B	806	-	4,4,4	0.92	0	6,6,6	0.45	0
8	PO4	C	802	-	4,4,4	0.92	0	6,6,6	0.41	0
8	PO4	B	802	-	4,4,4	0.91	0	6,6,6	0.44	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	C	801	PO4	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	699/728 (96%)	1.39	150 (21%) 0 0	37, 63, 208, 256	0
2	B	750/776 (96%)	0.36	34 (4%) 33 36	36, 61, 114, 153	0
3	C	730/797 (91%)	0.49	69 (9%) 8 8	40, 70, 118, 157	0
4	R	13/18 (72%)	-0.00	0 100 100	52, 62, 88, 90	0
5	V	16/16 (100%)	-0.27	1 (6%) 20 21	51, 54, 80, 141	0
6	X	5/28 (17%)	1.63	2 (40%) 0 0	66, 68, 77, 91	0
6	Y	9/28 (32%)	0.15	0 100 100	58, 66, 107, 119	0
All	All	2222/2391 (92%)	0.72	256 (11%) 4 4	36, 65, 175, 256	0

All (256) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	121	GLY	29.3
1	A	122	VAL	16.4
1	A	24	TYR	15.3
1	A	132	TYR	15.3
1	A	70	ALA	14.0
1	A	9	PHE	12.9
1	A	8	ASN	12.5
1	A	63	LYS	12.2
1	A	177	VAL	11.5
1	A	71	MET	10.9
1	A	47	MET	10.3
1	A	21	MET	10.1
1	A	1	MET	9.6
1	A	35	PHE	9.6
1	A	48	TYR	9.4
1	A	148	PHE	9.4
1	A	147	ILE	9.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	30	ASN	9.0
1	A	176	PHE	8.9
1	A	117	PHE	8.9
1	A	120	ILE	8.7
1	A	25	GLY	8.6
1	A	76	PHE	8.5
1	A	56	LEU	8.2
3	C	177	LEU	8.2
1	A	7	THR	8.2
1	A	155	MET	8.1
1	A	72	LEU	8.1
1	A	175	LEU	8.0
1	A	54	ILE	7.9
1	A	51	PHE	7.9
1	A	180	GLN	7.8
1	A	28	PRO	7.8
1	A	123	THR	7.8
1	A	145	ILE	7.6
1	A	66	ASP	7.6
1	A	124	ARG	7.5
1	A	64	GLU	7.4
3	C	349	LEU	7.4
1	A	178	LEU	7.3
3	C	352	LEU	7.2
1	A	67	ASP	7.2
1	A	45	CYS	7.2
1	A	40	THR	7.0
1	A	136	SER	7.0
1	A	104	ARG	7.0
3	C	420	PHE	6.8
1	A	183	ALA	6.8
3	C	495	ARG	6.8
1	A	52	HIS	6.8
1	A	13	ILE	6.6
3	C	418	LEU	6.6
1	A	62	VAL	6.5
2	B	647	PRO	6.4
1	A	128	GLU	6.4
1	A	146	HIS	6.4
1	A	163	LEU	6.3
1	A	130	TYR	6.3
3	C	354	MET	6.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	200	ILE	6.2
1	A	29	GLN	6.1
1	A	144	TYR	6.1
1	A	39	SER	6.1
2	B	734	ARG	6.1
1	A	126	LYS	6.0
1	A	27	ASN	6.0
1	A	173	THR	5.9
1	A	5	VAL	5.9
1	A	169	ALA	5.8
1	A	151	ASP	5.8
1	A	184	THR	5.8
1	A	131	TYR	5.8
1	A	182	LEU	5.7
1	A	140	GLY	5.7
1	A	11	PRO	5.7
1	A	164	ASP	5.6
1	A	152	GLY	5.6
3	C	494	ILE	5.6
2	B	203	ARG	5.6
1	A	4	PHE	5.5
1	A	57	GLU	5.5
1	A	75	ARG	5.4
1	A	162	ILE	5.4
1	A	127	VAL	5.3
1	A	171	ILE	5.3
2	B	187	LYS	5.1
1	A	26	GLU	5.1
3	C	351	THR	5.1
1	A	69	ASN	5.1
1	A	153	GLU	5.0
2	B	646	MET	5.0
3	C	348	ASN	4.9
2	B	650	GLY	4.9
1	A	185	ALA	4.9
1	A	125	ARG	4.8
1	A	65	ASN	4.8
1	A	42	MET	4.7
1	A	181	GLU	4.7
1	A	137	LYS	4.7
3	C	345	LEU	4.7
1	A	33	ASN	4.6

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	59	ASN	4.6
3	C	426	GLN	4.5
1	A	179	ARG	4.5
1	A	44	VAL	4.5
1	A	99	GLU	4.4
1	A	138	LEU	4.4
1	A	97	MET	4.4
1	A	12	MET	4.3
1	A	73	LYS	4.3
3	C	428	LEU	4.3
3	C	419	ASN	4.3
1	A	157	THR	4.3
1	A	34	LYS	4.2
2	B	197	LYS	4.2
2	B	652	ALA	4.2
1	A	129	ASP	4.2
1	A	154	GLU	4.2
1	A	22	LYS	4.2
3	C	713	LEU	4.2
1	A	143	VAL	4.2
2	B	193	ASP	4.2
1	A	112	TYR	4.2
3	C	686	ILE	4.1
3	C	607	GLU	4.1
1	A	84	ARG	4.1
1	A	31	GLU	4.1
1	A	10	ASN	4.1
1	A	38	ILE	4.1
2	B	199	MET	4.1
1	A	134	LYS	4.0
3	C	1	MET	4.0
1	A	3	ASN	4.0
2	B	188	LYS	4.0
1	A	20	THR	4.0
2	B	189	LYS	4.0
1	A	150	PHE	4.0
2	B	191	ILE	4.0
1	A	60	THR	3.9
3	C	421	VAL	3.9
1	A	53	PHE	3.8
3	C	738	LYS	3.7
1	A	114	THR	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	16	ARG	3.7
3	C	427	ARG	3.7
1	A	41	HIS	3.6
2	B	202	GLN	3.6
1	A	139	LYS	3.5
1	A	165	GLU	3.5
1	A	68	ASP	3.4
1	A	119	GLU	3.4
3	C	417	ASP	3.4
3	C	12	LYS	3.3
3	C	630	ARG	3.3
1	A	158	ASP	3.3
1	A	49	SER	3.3
3	C	716	LEU	3.3
2	B	201	THR	3.3
1	A	202	GLU	3.2
1	A	161	TYR	3.1
3	C	178	SER	3.1
1	A	14	LEU	3.1
3	C	444	THR	3.1
1	A	32	GLY	3.1
3	C	697	LEU	3.1
3	C	433	GLN	3.0
1	A	377	GLU	3.0
3	C	741	SER	3.0
2	B	198	LYS	3.0
1	A	6	ARG	3.0
3	C	707	VAL	2.9
3	C	721	LYS	2.8
3	C	350	GLN	2.8
1	A	549	ILE	2.8
3	C	628	LEU	2.8
2	B	741	SER	2.8
3	C	714	ASP	2.8
1	A	37	ALA	2.8
1	A	386	PHE	2.8
2	B	737	SER	2.8
3	C	685	SER	2.8
2	B	192	ARG	2.8
3	C	739	ARG	2.8
2	B	186	GLN	2.7
6	X	6	PRO	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	172	LYS	2.7
3	C	346	THR	2.7
2	B	736	SER	2.7
3	C	347	GLY	2.7
1	A	383	PRO	2.7
1	A	201	LEU	2.6
2	B	633	HIS	2.6
2	B	733	GLY	2.6
1	A	79	ILE	2.6
1	A	714	GLY	2.6
3	C	403	VAL	2.6
3	C	740	ASP	2.6
1	A	36	ALA	2.6
3	C	344	VAL	2.5
1	A	118	ILE	2.5
3	C	434	LEU	2.5
3	C	175	ARG	2.5
3	C	724	VAL	2.5
2	B	732	SER	2.5
3	C	321	ASN	2.5
1	A	101	SER	2.5
3	C	299	ARG	2.5
3	C	356	ILE	2.5
1	A	135	ALA	2.4
3	C	180	ASP	2.4
3	C	5	LYS	2.4
3	C	699	LYS	2.4
1	A	106	LEU	2.4
1	A	109	LEU	2.4
1	A	174	ARG	2.4
3	C	318	ARG	2.4
3	C	701	ASN	2.3
2	B	190	ARG	2.3
3	C	447	LYS	2.3
3	C	353	THR	2.3
3	C	440	LYS	2.3
1	A	58	GLY	2.3
3	C	183	LEU	2.3
3	C	8	MET	2.3
1	A	170	ARG	2.3
1	A	167	SER	2.3
2	B	154	GLY	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
5	V	16	G	2.3
1	A	141	GLU	2.2
2	B	735	ILE	2.2
3	C	430	PRO	2.2
3	C	742	SER	2.2
2	B	706	ARG	2.2
3	C	532	ASN	2.2
1	A	546	ARG	2.2
3	C	703	LYS	2.1
3	C	285	HIS	2.1
3	C	170	ASN	2.1
1	A	15	GLU	2.1
3	C	507	GLU	2.1
1	A	149	SER	2.1
2	B	750	ILE	2.1
2	B	730	LEU	2.1
3	C	733	LEU	2.1
1	A	23	GLU	2.1
1	A	378	ASP	2.1
2	B	185	PHE	2.1
3	C	153	ASP	2.1
2	B	208	LYS	2.1
3	C	152	LYS	2.1
1	A	102	LYS	2.1
1	A	160	GLU	2.0
2	B	740	PHE	2.0
1	A	2	GLU	2.0
3	C	562	THR	2.0
6	X	2	SER	2.0

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

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
6	SEP	X	5	10/11	0.93	0.17	65,74,89,92	0
6	SEP	Y	19	10/11	0.97	0.14	52,57,63,63	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( $\text{\AA}^2$ )	Q<0.9
8	PO4	C	804	5/5	0.81	0.33	75,83,94,98	0
8	PO4	A	802	5/5	0.82	0.25	93,93,110,110	0
8	PO4	B	807	5/5	0.86	0.33	78,81,84,105	0
8	PO4	C	805	5/5	0.88	0.20	88,93,95,98	0
8	PO4	B	804	5/5	0.88	0.19	47,59,67,89	0
8	PO4	C	801	5/5	0.88	0.15	72,83,90,100	0
8	PO4	A	803	5/5	0.89	0.23	67,74,91,92	0
8	PO4	A	804	5/5	0.89	0.23	94,101,119,125	0
8	PO4	B	806	5/5	0.92	0.14	95,105,116,120	0
8	PO4	C	803	5/5	0.93	0.14	62,67,81,86	0
8	PO4	B	802	5/5	0.93	0.17	68,69,85,92	0
8	PO4	B	803	5/5	0.94	0.12	65,68,80,83	0
8	PO4	C	802	5/5	0.95	0.11	75,77,93,93	0
8	PO4	B	801	5/5	0.97	0.11	64,67,93,94	0
7	ZN	A	801	1/1	0.97	0.22	74,74,74,74	0
8	PO4	B	805	5/5	0.99	0.11	63,66,67,72	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.