



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 29, 2024 – 12:25 am GMT

PDB ID : 6FHH  
Title : Crystal structure of bat influenza A/H17N10 polymerase with viral RNA promoter bound to a 22-mer modified Pol II CTD peptide with serine 5 thiophosphorylated.  
Authors : Lukarska, M.; Cusack, S.  
Deposited on : 2018-01-14  
Resolution : 2.70 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

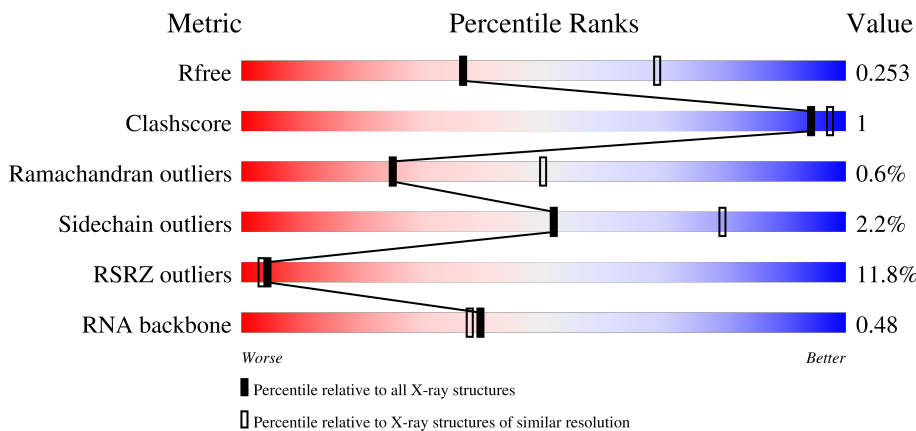
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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.



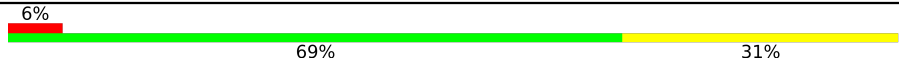


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)
RNA backbone	3102	1159 (3.00-2.40)

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 of 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	738	
2	B	776	
3	C	809	
4	R	18	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	V	16	 6% 69% 31%
6	X	22	 27% 5% 68%
6	Y	22	 41% 5% 55%

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 18457 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
			Total	C	N	O	S			
1	A	709	5770	3665	971	1097	37	0	0	0

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	GLY	-	expression tag	UNP H6QM92
A	-12	SER	-	expression tag	UNP H6QM92
A	-11	HIS	-	expression tag	UNP H6QM92
A	-10	HIS	-	expression tag	UNP H6QM92
A	-9	HIS	-	expression tag	UNP H6QM92
A	-8	HIS	-	expression tag	UNP H6QM92
A	-7	HIS	-	expression tag	UNP H6QM92
A	-6	HIS	-	expression tag	UNP H6QM92
A	-5	HIS	-	expression tag	UNP H6QM92
A	-4	HIS	-	expression tag	UNP H6QM92
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
			Total	C	N	O	S			
2	B	750	5986	3766	1062	1118	40	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	GLY	-	expression tag	UNP H6QM91
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
			Total	C	N	O	S			
3	C	725	5753	3631	1015	1075	32	0	0	0

There are 49 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

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
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
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	GLY	-	expression tag	UNP H6QM90
C	772	GLY	-	expression tag	UNP H6QM90
C	773	SER	-	expression tag	UNP H6QM90
C	774	GLY	-	expression tag	UNP H6QM90
C	775	GLY	-	expression tag	UNP H6QM90
C	776	GLY	-	expression tag	UNP H6QM90
C	777	SER	-	expression tag	UNP H6QM90
C	778	GLY	-	expression tag	UNP H6QM90
C	779	GLY	-	expression tag	UNP H6QM90
C	780	SER	-	expression tag	UNP H6QM90
C	781	ALA	-	expression tag	UNP H6QM90
C	782	TRP	-	expression tag	UNP H6QM90
C	783	SER	-	expression tag	UNP H6QM90
C	784	HIS	-	expression tag	UNP H6QM90
C	785	PRO	-	expression tag	UNP H6QM90
C	786	GLN	-	expression tag	UNP H6QM90
C	787	PHE	-	expression tag	UNP H6QM90
C	788	GLU	-	expression tag	UNP H6QM90
C	789	LYS	-	expression tag	UNP H6QM90
C	790	GLY	-	expression tag	UNP H6QM90
C	791	ARG	-	expression tag	UNP H6QM90
C	792	SER	-	expression tag	UNP H6QM90
C	793	GLY	-	expression tag	UNP H6QM90
C	794	GLY	-	expression tag	UNP H6QM90
C	795	GLU	-	expression tag	UNP H6QM90
C	796	ASN	-	expression tag	UNP H6QM90
C	797	LEU	-	expression tag	UNP H6QM90
C	798	TYR	-	expression tag	UNP H6QM90
C	799	PHE	-	expression tag	UNP H6QM90
C	800	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
			Total	C	N	O	P			
4	R	13	247	111	37	87	12	0	0	0

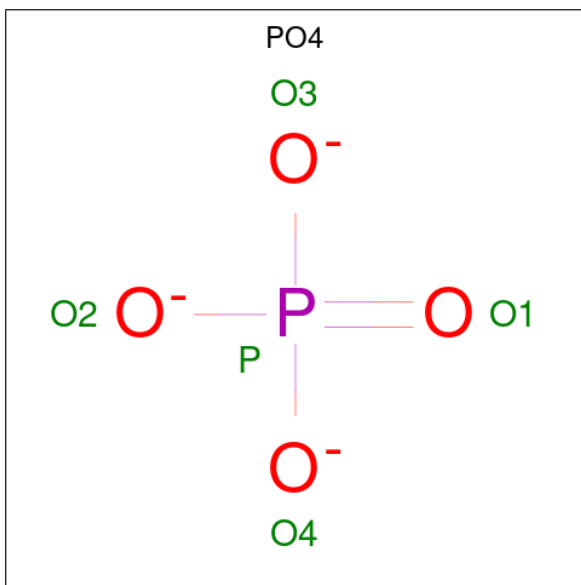
- Molecule 5 is a RNA chain called RNA (5'-R(P\*AP\*GP\*UP\*AP\*GP\*UP\*AP\*AP\*CP\*AP\*AP\*GP\*AP\*GP\*GP\*G)-3').

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

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
6	X	7	55	32	7	14	1	1	0	0	0
6	Y	10	80	49	10	19	1	1	0	0	0

- Molecule 7 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	1	Total	Zn	0	0
			1	1		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	41	Total	O	0	0
			41	41		
9	B	54	Total	O	0	0
			54	54		

*Continued on next page...*



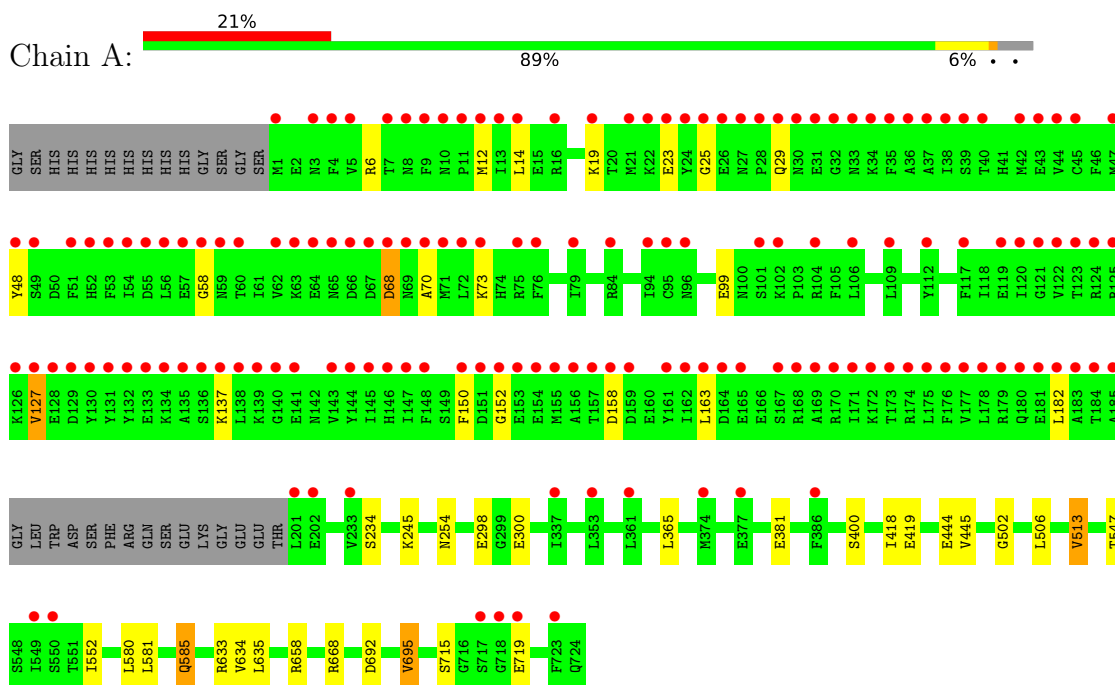
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
9	C	35	Total 35	O 35	0	0
9	R	1	Total 1	O 1	0	0
9	V	6	Total 6	O 6	0	0

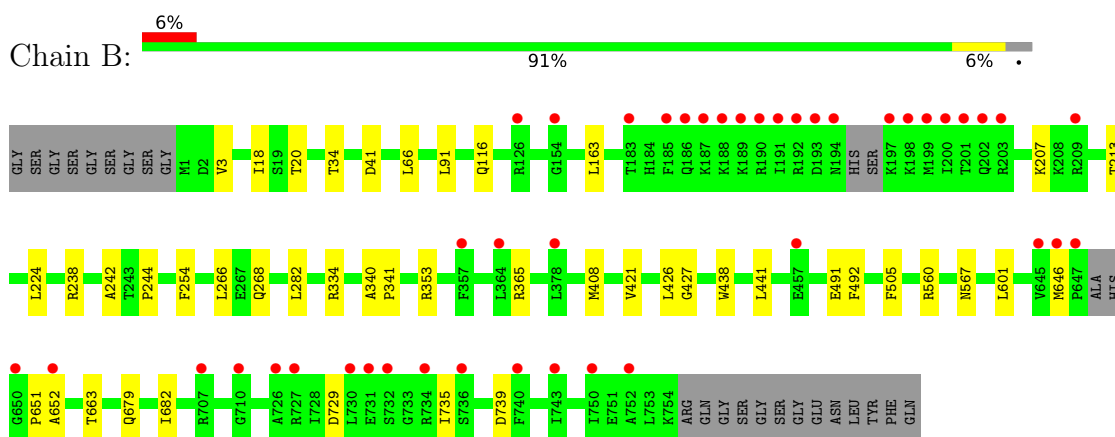
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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, $\alpha$ , $\beta$ , $\gamma$	267.77Å 147.81Å 88.41Å 90.00° 97.08° 90.00°	Depositor
Resolution (Å)	50.00 – 2.70 49.41 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.6 (50.00-2.70) 98.7 (49.41-2.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.45 (at 2.69Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.213 , 0.258 0.214 , 0.253	Depositor DCC
$R_{free}$ test set	4534 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	67.9	Xtrriage
Anisotropy	0.333	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 51.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	18457	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	92.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, ZN, 2RX

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	0.38	0/5892	0.57	0/7934
2	B	0.37	0/6100	0.60	0/8229
3	C	0.36	0/5851	0.58	0/7897
4	R	0.26	0/273	0.68	0/421
5	V	0.56	1/397 (0.3%)	0.70	0/617
6	X	0.46	0/46	0.41	0/61
6	Y	0.34	0/73	0.47	0/98
All	All	0.37	1/18632 (0.0%)	0.59	0/25257

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	V	1	A	OP3-P	-10.16	1.49	1.61

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	5770	0	5678	19	0
2	B	5986	0	6006	22	0
3	C	5753	0	5876	14	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	R	247	0	128	0	0
5	V	353	0	175	0	0
6	X	55	0	47	0	0
6	Y	80	0	66	0	0
7	A	15	0	0	0	0
7	B	35	0	0	0	0
7	C	25	0	0	1	0
8	B	1	0	0	0	0
9	A	41	0	0	0	0
9	B	54	0	0	0	0
9	C	35	0	0	0	0
9	R	1	0	0	0	0
9	V	6	0	0	0	0
All	All	18457	0	17976	48	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 1.

All (48) 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:719:GLU:OE1	1:A:719:GLU:N	2.20	0.72
1:A:506:LEU:HD11	1:A:513:VAL:HG22	1.73	0.69
1:A:719:GLU:H	1:A:719:GLU:CD	1.99	0.66
1:A:445:VAL:CG1	1:A:580:LEU:HD22	2.26	0.65
1:A:585:GLN:HG2	2:B:505:PHE:CD2	2.37	0.59
3:C:145:VAL:HG21	3:C:220:ILE:HD11	1.85	0.59
1:A:445:VAL:HG13	1:A:580:LEU:HD22	1.85	0.57
2:B:224:LEU:HD22	2:B:244:PRO:HA	1.85	0.57
1:A:547:THR:HG21	1:A:552:ILE:HD12	1.89	0.54
2:B:34:THR:HG22	2:B:353:ARG:HB2	1.90	0.52
3:C:395:LEU:HD21	3:C:463:ILE:HD11	1.92	0.51
3:C:554:ILE:HD12	3:C:672:VAL:HG13	1.92	0.51
2:B:282:LEU:HG	2:B:441:LEU:HD22	1.92	0.51
3:C:144:ARG:NH2	7:C:901:PO4:O2	2.44	0.50
3:C:145:VAL:HG11	3:C:220:ILE:CD1	2.42	0.50
3:C:554:ILE:HD11	3:C:665:ILE:HD12	1.94	0.49
3:C:462:GLY:HA3	3:C:474:THR:HB	1.93	0.49
2:B:427:GLY:HA3	2:B:438:TRP:CD1	2.48	0.48
1:A:506:LEU:HD11	1:A:513:VAL:CG2	2.43	0.48
2:B:679:GLN:HB3	2:B:682:ILE:HD12	1.96	0.48

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:242:ALA:HB2	2:B:408:MET:HE3	1.95	0.47
1:A:635:LEU:HD22	2:B:20:THR:HG22	1.96	0.47
3:C:114:VAL:HG22	3:C:193:LEU:HD22	1.96	0.47
2:B:163:LEU:HD23	2:B:254:PHE:CZ	2.50	0.46
2:B:340:ALA:HB3	2:B:341:PRO:HD3	1.97	0.46
3:C:696:ILE:HG23	3:C:732:VAL:HG13	1.98	0.46
2:B:34:THR:HG22	2:B:353:ARG:CB	2.46	0.45
3:C:395:LEU:HD21	3:C:463:ILE:CD1	2.47	0.45
2:B:91:LEU:HD23	2:B:426:LEU:HD13	1.98	0.44
2:B:441:LEU:HD23	2:B:492:PHE:CZ	2.52	0.44
1:A:444:GLU:OE2	1:A:633:ARG:NH1	2.51	0.44
2:B:242:ALA:HB2	2:B:408:MET:CE	2.48	0.43
1:A:300:GLU:HG3	2:B:567:ASN:HB2	2.01	0.43
1:A:418:ILE:HD13	1:A:580:LEU:HD11	2.01	0.43
2:B:268:GLN:HB2	2:B:421:VAL:HG13	2.01	0.43
3:C:138:GLN:HE21	3:C:221:ALA:HB1	1.84	0.43
1:A:6:ARG:HA	1:A:14:LEU:HD21	2.01	0.42
1:A:658:ARG:HD3	2:B:18:ILE:HD11	2.01	0.42
2:B:266:LEU:HD13	2:B:421:VAL:HG11	2.02	0.42
2:B:663:THR:HG21	3:C:99:TRP:CD1	2.55	0.41
2:B:729:ASP:HB3	2:B:735:ILE:HD12	2.02	0.41
1:A:445:VAL:HG11	1:A:580:LEU:HD22	2.00	0.41
3:C:696:ILE:HG23	3:C:732:VAL:CG1	2.51	0.41
1:A:365:LEU:HD13	1:A:502:GLY:HA2	2.02	0.41
3:C:167:VAL:HG23	3:C:168:PHE:CD2	2.55	0.41
1:A:634:VAL:HG22	2:B:3:VAL:HG11	2.03	0.41
1:A:692:ASP:HB3	1:A:695:VAL:HG12	2.03	0.41
1:A:234:SER:OG	2:B:334:ARG:NH2	2.50	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	705/738 (96%)	652 (92%)	43 (6%)	10 (1%)	11	28
2	B	744/776 (96%)	718 (96%)	24 (3%)	2 (0%)	41	66
3	C	719/809 (89%)	692 (96%)	27 (4%)	0	100	100
6	X	4/22 (18%)	3 (75%)	1 (25%)	0	100	100
6	Y	7/22 (32%)	7 (100%)	0	0	100	100
All	All	2179/2367 (92%)	2072 (95%)	95 (4%)	12 (1%)	25	50

All (12) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	19	LYS
1	A	68	ASP
1	A	127	VAL
1	A	137	LYS
2	B	651	PRO
1	A	58	GLY
1	A	23	GLU
1	A	70	ALA
1	A	25	GLY
1	A	99	GLU
2	B	652	ALA
1	A	152	GLY

### 5.3.2 Protein sidechains [i](#)

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	633/657 (96%)	611 (96%)	22 (4%)	36	65
2	B	659/676 (98%)	647 (98%)	12 (2%)	59	83
3	C	642/706 (91%)	634 (99%)	8 (1%)	71	88
6	X	6/19 (32%)	6 (100%)	0	100	100
6	Y	9/19 (47%)	9 (100%)	0	100	100
All	All	1949/2077 (94%)	1907 (98%)	42 (2%)	52	79



All (42) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	MET
1	A	29	GLN
1	A	48	TYR
1	A	68	ASP
1	A	73	LYS
1	A	127	VAL
1	A	150	PHE
1	A	158	ASP
1	A	163	LEU
1	A	182	LEU
1	A	245	LYS
1	A	254	ASN
1	A	298	GLU
1	A	381	GLU
1	A	400	SER
1	A	419	GLU
1	A	513	VAL
1	A	581	LEU
1	A	585	GLN
1	A	668	ARG
1	A	695	VAL
1	A	715	SER
2	B	41	ASP
2	B	66	LEU
2	B	116	GLN
2	B	207	LYS
2	B	213	THR
2	B	238	ARG
2	B	365	ARG
2	B	491	GLU
2	B	560	ARG
2	B	601	LEU
2	B	646	MET
2	B	739	ASP
3	C	15	ARG
3	C	164	MET
3	C	354	MET
3	C	410	MET
3	C	536	MET
3	C	627	SER
3	C	672	VAL
3	C	678	SER

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

Mol	Chain	Res	Type
1	A	30	ASN
2	B	313	GLN
3	C	379	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	R	11/18 (61%)	1 (9%)	0
5	V	15/16 (93%)	3 (20%)	1 (6%)
All	All	26/34 (76%)	4 (15%)	1 (3%)

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
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
5	V	5	G

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

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	2RX	Y	19	6	6,9,10	1.54	2 (33%)	4,12,14	1.22	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	2RX	X	5	6	6,9,10	1.27	2 (33%)	4,12,14	2.04	2 (50%)

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	2RX	Y	19	6	-	0/4/8/10	-
6	2RX	X	5	6	-	1/4/8/10	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	Y	19	2RX	P40-OG	2.69	1.62	1.57
6	Y	19	2RX	P40-O42	-2.34	1.50	1.56
6	X	5	2RX	P40-O42	-2.04	1.51	1.56
6	X	5	2RX	P40-OG	2.02	1.61	1.57

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	X	5	2RX	OG-CB-CA	3.03	111.09	108.14
6	X	5	2RX	OG-P40-O43	-2.38	105.65	114.42

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	X	5	2RX	N-CA-CB-OG

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides 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
7	PO4	C	904	-	4,4,4	1.01	0	6,6,6	0.40	0
7	PO4	C	902	-	4,4,4	0.96	0	6,6,6	0.46	0
7	PO4	A	803	-	4,4,4	0.89	0	6,6,6	0.39	0
7	PO4	C	905	-	4,4,4	0.95	0	6,6,6	0.41	0
7	PO4	B	806	-	4,4,4	0.97	0	6,6,6	0.55	0
7	PO4	B	802	-	4,4,4	0.83	0	6,6,6	0.51	0
7	PO4	C	903	-	4,4,4	0.86	0	6,6,6	0.42	0
7	PO4	C	901	-	4,4,4	0.94	0	6,6,6	0.36	0
7	PO4	B	805	-	4,4,4	0.82	0	6,6,6	0.52	0
7	PO4	B	804	-	4,4,4	0.82	0	6,6,6	0.48	0
7	PO4	B	808	-	4,4,4	0.89	0	6,6,6	0.43	0
7	PO4	A	802	-	4,4,4	0.91	0	6,6,6	0.56	0
7	PO4	A	801	-	4,4,4	0.89	0	6,6,6	0.50	0
7	PO4	B	803	-	4,4,4	0.85	0	6,6,6	0.50	0
7	PO4	B	807	-	4,4,4	0.93	0	6,6,6	0.52	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
7	C	901	PO4	1	0

## 5.7 Other polymers [i](#)

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	709/738 (96%)	1.40	155 (21%) 0 0	49, 80, 228, 288	0
2	B	750/776 (96%)	0.45	43 (5%) 23 22	45, 77, 136, 177	0
3	C	725/809 (89%)	0.54	65 (8%) 9 7	49, 83, 132, 162	0
4	R	13/18 (72%)	0.18	0 100 100	62, 74, 111, 113	0
5	V	16/16 (100%)	-0.21	1 (6%) 20 19	64, 70, 103, 167	0
6	X	6/22 (27%)	0.15	0 100 100	70, 74, 98, 110	0
6	Y	9/22 (40%)	0.64	0 100 100	61, 66, 108, 125	0
All	All	2228/2401 (92%)	0.77	264 (11%) 4 3	45, 80, 193, 288	0

All (264) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	121	GLY	23.9
1	A	8	ASN	15.1
1	A	132	TYR	14.8
1	A	71	MET	14.7
1	A	177	VAL	12.5
1	A	24	TYR	12.3
1	A	54	ILE	12.0
1	A	9	PHE	11.8
1	A	48	TYR	11.1
1	A	63	LYS	11.1
1	A	162	ILE	10.9
1	A	1	MET	10.5
1	A	42	MET	10.3
1	A	122	VAL	10.0
1	A	148	PHE	10.0
1	A	47	MET	10.0
1	A	163	LEU	9.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	45	CYS	9.4
1	A	64	GLU	9.4
1	A	62	VAL	9.3
1	A	35	PHE	9.1
1	A	117	PHE	9.0
1	A	7	THR	8.6
1	A	4	PHE	8.6
1	A	136	SER	8.5
1	A	13	ILE	8.4
1	A	178	LEU	8.3
1	A	180	GLN	8.3
1	A	49	SER	8.3
1	A	144	TYR	8.3
3	C	418	LEU	8.3
1	A	28	PRO	8.2
1	A	51	PHE	7.9
1	A	70	ALA	7.8
1	A	145	ILE	7.8
1	A	175	LEU	7.7
1	A	184	THR	7.5
3	C	349	LEU	7.3
3	C	177	LEU	7.3
2	B	199	MET	7.1
1	A	158	ASP	7.0
1	A	171	ILE	7.0
1	A	66	ASP	6.9
1	A	57	GLU	6.9
1	A	176	PHE	6.7
1	A	5	VAL	6.7
3	C	354	MET	6.7
1	A	52	HIS	6.6
1	A	155	MET	6.5
1	A	183	ALA	6.4
1	A	146	HIS	6.3
1	A	73	LYS	6.2
3	C	345	LEU	6.1
1	A	173	THR	6.1
2	B	200	ILE	6.1
1	A	21	MET	6.1
3	C	420	PHE	6.0
2	B	203	ARG	6.0
1	A	22	LYS	5.9

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	34	LYS	5.9
1	A	124	ARG	5.9
1	A	120	ILE	5.9
1	A	151	ASP	5.8
1	A	131	TYR	5.8
1	A	123	THR	5.8
1	A	25	GLY	5.7
3	C	713	LEU	5.7
2	B	650	GLY	5.7
1	A	72	LEU	5.7
1	A	147	ILE	5.7
2	B	187	LYS	5.6
1	A	104	ARG	5.5
3	C	419	ASN	5.5
3	C	352	LEU	5.5
1	A	40	THR	5.4
2	B	188	LYS	5.4
1	A	59	ASN	5.4
1	A	75	ARG	5.3
1	A	126	LYS	5.3
1	A	185	ALA	5.2
2	B	647	PRO	5.2
3	C	431	MET	5.2
1	A	11	PRO	5.1
3	C	738	LYS	5.1
1	A	67	ASP	5.1
2	B	198	LYS	5.1
2	B	202	GLN	5.1
1	A	723	PHE	4.9
1	A	125	ARG	4.9
2	B	734	ARG	4.9
3	C	1	MET	4.9
3	C	495	ARG	4.8
1	A	33	ASN	4.8
2	B	186	GLN	4.8
1	A	134	LYS	4.7
2	B	193	ASP	4.7
1	A	60	THR	4.7
1	A	76	PHE	4.7
3	C	686	ILE	4.7
2	B	201	THR	4.6
1	A	182	LEU	4.6

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	157	THR	4.6
1	A	169	ALA	4.6
1	A	179	ARG	4.6
1	A	127	VAL	4.5
1	A	56	LEU	4.5
1	A	181	GLU	4.5
1	A	27	ASN	4.5
1	A	44	VAL	4.4
1	A	38	ILE	4.4
1	A	128	GLU	4.4
1	A	130	TYR	4.4
2	B	652	ALA	4.4
2	B	646	MET	4.4
1	A	156	ALA	4.4
3	C	346	THR	4.3
1	A	164	ASP	4.3
3	C	434	LEU	4.2
3	C	716	LEU	4.2
1	A	139	LYS	4.1
1	A	112	TYR	4.1
1	A	165	GLU	4.1
3	C	697	LEU	4.0
2	B	185	PHE	4.0
1	A	154	GLU	4.0
5	V	16	G	4.0
1	A	37	ALA	3.9
2	B	189	LYS	3.9
3	C	699	LYS	3.9
2	B	191	ILE	3.9
3	C	348	ASN	3.9
1	A	3	ASN	3.8
1	A	14	LEU	3.8
2	B	194	ASN	3.7
1	A	32	GLY	3.7
1	A	39	SER	3.7
1	A	16	ARG	3.7
3	C	185	ILE	3.7
1	A	36	ALA	3.6
1	A	12	MET	3.6
3	C	353	THR	3.6
1	A	26	GLU	3.5
1	A	29	GLN	3.5

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
3	C	351	THR	3.5
1	A	138	LEU	3.5
1	A	19	LYS	3.5
1	A	53	PHE	3.5
1	A	55	ASP	3.5
1	A	79	ILE	3.5
1	A	137	LYS	3.4
2	B	197	LYS	3.4
3	C	417	ASP	3.4
2	B	730	LEU	3.4
1	A	10	ASN	3.4
1	A	129	ASP	3.3
1	A	65	ASN	3.3
1	A	377	GLU	3.3
1	A	153	GLU	3.3
1	A	102	LYS	3.3
2	B	192	ARG	3.3
2	B	731	GLU	3.2
3	C	739	ARG	3.2
3	C	685	SER	3.1
2	B	740	PHE	3.1
1	A	152	GLY	3.1
1	A	718	GLY	3.1
3	C	725	LEU	3.1
2	B	743	ILE	3.1
1	A	135	ALA	3.0
1	A	30	ASN	3.0
3	C	715	LYS	2.9
1	A	69	ASN	2.9
3	C	740	ASP	2.9
2	B	154	GLY	2.9
1	A	159	ASP	2.9
3	C	607	GLU	2.9
1	A	84	ARG	2.8
1	A	717	SER	2.8
3	C	433	GLN	2.8
3	C	694	PHE	2.8
3	C	437	HIS	2.8
1	A	201	LEU	2.8
1	A	172	LYS	2.8
2	B	378	LEU	2.7
2	B	190	ARG	2.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	B	357	PHE	2.7
2	B	750	ILE	2.7
1	A	143	VAL	2.7
1	A	133	GLU	2.7
3	C	180	ASP	2.7
1	A	23	GLU	2.7
3	C	722	ALA	2.7
2	B	126	ARG	2.7
1	A	119	GLU	2.7
1	A	170	ARG	2.7
2	B	732	SER	2.7
3	C	347	GLY	2.6
1	A	109	LEU	2.6
1	A	68	ASP	2.6
1	A	141	GLU	2.6
1	A	337	ILE	2.6
1	A	549	ILE	2.6
3	C	299	ARG	2.6
3	C	724	VAL	2.6
1	A	150	PHE	2.6
1	A	374	MET	2.6
1	A	386	PHE	2.5
1	A	31	GLU	2.5
3	C	741	SER	2.5
1	A	719	GLU	2.5
3	C	566	THR	2.5
1	A	550	SER	2.5
2	B	710	GLY	2.4
3	C	446	LEU	2.4
2	B	457	GLU	2.4
3	C	438	PHE	2.4
3	C	693	GLY	2.4
1	A	168	ARG	2.4
1	A	361	LEU	2.4
1	A	101	SER	2.4
1	A	96	ASN	2.4
3	C	630	ARG	2.4
1	A	140	GLY	2.3
3	C	12	LYS	2.3
1	A	43	GLU	2.3
2	B	645	VAL	2.3
3	C	154	LEU	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	58	GLY	2.3
2	B	726	ALA	2.3
1	A	94	ILE	2.3
3	C	430	PRO	2.3
3	C	13	ASN	2.3
1	A	95	CYS	2.3
3	C	714	ASP	2.3
2	B	364	LEU	2.2
3	C	701	ASN	2.2
1	A	174	ARG	2.2
1	A	161	TYR	2.2
3	C	367	GLY	2.2
1	A	233	VAL	2.2
3	C	413	ALA	2.2
1	A	167	SER	2.2
2	B	727	ARG	2.2
3	C	721	LYS	2.2
1	A	353	LEU	2.1
2	B	209	ARG	2.1
3	C	706	PRO	2.1
2	B	752	ALA	2.1
3	C	344	VAL	2.1
3	C	2	GLU	2.1
3	C	735	MET	2.1
1	A	202	GLU	2.1
3	C	494	ILE	2.1
3	C	700	ALA	2.1
3	C	321	ASN	2.1
3	C	677	ARG	2.1
3	C	414	VAL	2.0
2	B	183	THR	2.0
3	C	447	LYS	2.0
2	B	736	SER	2.0
3	C	565	MET	2.0
3	C	444	THR	2.0
1	A	106	LEU	2.0
2	B	707	ARG	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(Å <sup>2</sup> )	Q<0.9
6	2RX	X	5	10/11	0.96	0.11	68,78,89,94	0
6	2RX	Y	19	10/11	0.97	0.20	61,64,66,67	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
7	PO4	A	801	5/5	0.74	0.29	140,144,146,148	0
7	PO4	B	805	5/5	0.81	0.21	101,106,113,115	0
7	PO4	C	904	5/5	0.83	0.23	108,110,113,113	0
7	PO4	B	808	5/5	0.87	0.33	110,111,116,119	0
7	PO4	B	802	5/5	0.87	0.16	90,90,101,101	0
7	PO4	C	901	5/5	0.88	0.16	99,99,105,107	0
7	PO4	B	806	5/5	0.88	0.24	95,101,111,114	0
7	PO4	A	802	5/5	0.90	0.22	88,92,98,102	0
7	PO4	A	803	5/5	0.90	0.19	126,129,133,133	0
7	PO4	C	905	5/5	0.90	0.23	123,126,128,132	0
7	PO4	B	804	5/5	0.92	0.18	98,100,105,109	0
7	PO4	B	803	5/5	0.93	0.17	98,100,101,107	0
7	PO4	C	903	5/5	0.93	0.12	98,101,108,109	0
7	PO4	B	807	5/5	0.96	0.14	89,90,93,93	0
7	PO4	C	902	5/5	0.97	0.17	95,97,98,99	0
8	ZN	B	801	1/1	0.99	0.19	82,82,82,82	0

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

There are no such residues in this entry.